Advances in Peer-Led Learning

Number 4 | Fall 2024 | Article 7

Testing the Reliability of ChatGPT in Providing Spectral Information of Organic Molecules

Ana Fraiman*
Department of Chemistry
Northeastern Illinois University
Chicago, Illinois 60625
a-fraiman@neiu.edu

Paola Barron, Brooke Dorsey
University of Texas at El Paso

Sofia A. Delgado
Yale University

Emanuel Sanchez
University of Texas Rio Grande Valley

Recommended Citation

Fraiman, A., Barron, P., Dorsey, B., Delgado, S.A., & Sanchez, E. (2024). Testing the reliability of ChatGPT in providing spectral information of organic molecules. *Advances in Peer-Led Learning, 4,* 69-108. Online at https://doi.org/10.54935/apll2024-01-07-69

Advances in Peer-Led Learning
2024, Number 4, 69-108

Testing the Reliability of ChatGPT in Providing Spectral Information

Ana Fraiman*, Paola Barron, Brooke Dorsey, Sofia A. Delgado, & Emanuel Sanchez

Department of Chemistry
Northeastern Illinois University
Chicago, Illinois 60625
a-fraiman@neiu.edu

Abstract

ChatGPT has emerged as a tool for predicting Hydrogen-1 Nuclear Magnetic Resonance (1HNMR), Infrared (IR), and Mass Spectrometry (MS) spectra. In order to rely on ChatGPT as a tool, a proof of concept for its application was essential. Having the ChatGPT outputs be validated by experimental data and databases (SDBS) has proven valuable for future uses. By inference, ChatGPT spectra output can be used to design problems without interference of copyright infringement of existing databases.

Keywords: Internet/Web-Based Learning, Generative Artificial Intelligence, Organic Chemistry, HNMR Spectroscopy, IR Spectroscopy, Mass Spectrometry

<u>The development of ChatGPT</u>

Technology is rapidly advancing in education and expanding across various domains, encompassing cell phones, smartwatches, computers, and nearly all digital technologies. This progress has played a pivotal role in driving innovation for an Organic Chemistry II workbook, largely attributed to the contributions of the Chat Generative Pre-Trained Transformer (ChatGPT). ChatGPT is an artificial intelligence (AI) language model developed by OpenAI, a prominent AI research company (OpenAI, 2022). It made its debut on November 30, 2022, and has swiftly gained popularity due to its accessibility and prompt responses.

ChatGPT serves as a valuable resource for addressing a wide range of inquiries posed (OpenAI, 2022). It excels in providing information on specific topics, assisting with essay writing, and even aiding in coding tasks. We address here how the outputs for spectroscopy can be provided by ChatGPT.

<u>Previous applications of ChatGPT in chemistry</u>

Since the introduction of ChatGPT, a variety of applications for artificial intelligence have emerged. Recently, these studies have looked at applications in higher education. With the integration of technology in the classroom increasing rapidly over the past few years, it is only natural that generative AI becomes a tool for students (Leon & Vidhani, 2023). AI can be used as an aid to help a student further understand or check their assignment for errors (Castro Nascimento & Pimentel, 2023). Perhaps the discipline most obviously impacted by the emergence of ChatGPT is computer science, as students and professionals alike can refer to it as a resource for writing code more efficiently, as well as cybersecurity (Hörnemalm, 2023). The use of ChatGPT as a "discussion partner" when completing coursework could be helpful in any field (Malinka et al., 2023). ChatGPT has shown it can be useful in helping students perform by explaining some calculations and providing assistance in written chemistry laboratory reports (Humphry & Fuller, 2023).

Although AI proves to have many benefits when incorporated into higher education, there are also concerns about its usage. Namely, there is an ethical concern of academic dishonesty from students who may take advantage of the ease with which they can obtain assistance from ChatGPT (Rawas, 2024; Terry, 2023). In educational matters, ChatGPT has been seen as a supplemental learning tool for students to assist them in learning general chemistry concepts as well as in other disciplines (Emenike & Emenike, 2023; Rawas, 2024). Previous work has used ChatGPT to develop practice problems for general chemistry (Humphry & Fuller, 2023; Christiansen, 2024).

<u>Testing ChatGPT for spectroscopy practice problems</u>

Preliminary data showed that ChatGPT could correctly predict spectra characteristics for some molecules. The aim of this project is to authenticate ChatGPT's capability to provide the same information for various molecules and ascertain whether its data output aligned with experimentally collected spectra.

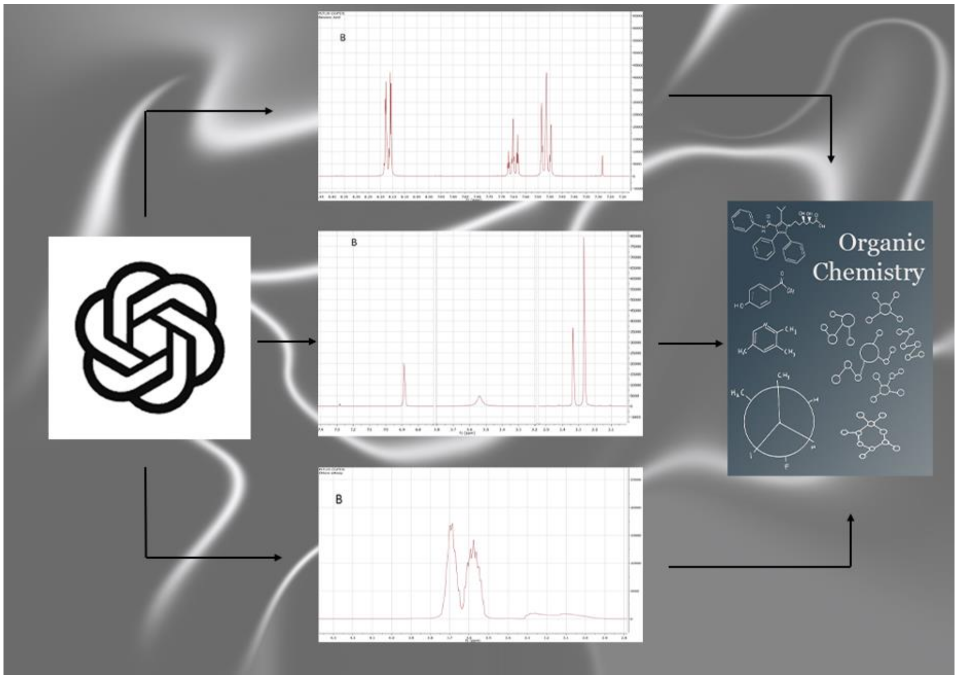
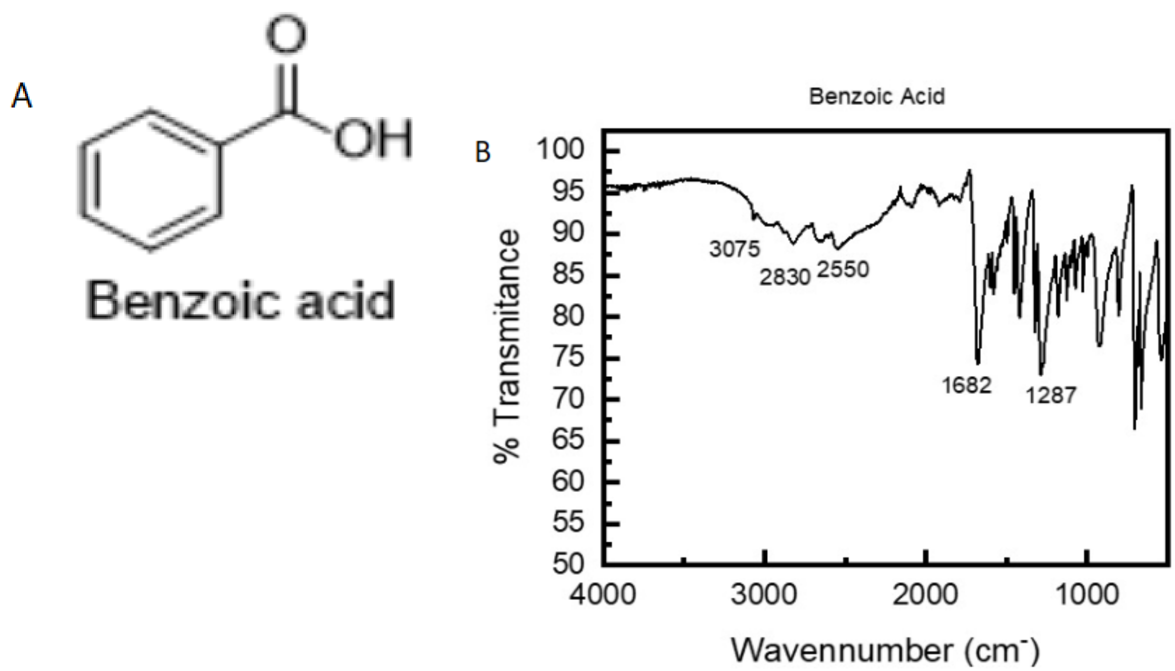


Figure 1. ChatGPT aids Spectroscopy. Visual representation of the use of ChatGPT to provide textual information of various spectra for use in the creation of the PLTLIS Organic Chemistry II workbook. The spectra shown above were experimentally collected to validate the textual output provided by ChatGPT that can be used for the new workbook.

To test the capabilities and limits of the free version of ChatGPT, the AI model was prompted to provide spectra data of target molecules in an attempt to validate its information. The spectral information will be used for the construction of spectroscopy practice problems for a second semester Organic Chemistry workbook without the infringement of copyright spectral data. This new Organic Chemistry II workbook will be a sequel to the Organic Chemistry I workbook (Fioramonti, et al., 2018).

ChatGPT proves to be a tool for spectral information. In Infrared (IR) and Mass Spectrometry (MS) ChatGPT provides the prominent peaks for the compound of interest and guides how to interpret these spectra for eleven compounds (see Appendix, Figures 1-33). ChatGPT provides a numerical description of prominent peaks that were expected in the spectra and explains what parts of the chemical structure were responsible for these peaks and why. For Nuclear Magnetic Resonance (NMR), ChatGPT provides the expected chemical shifts and explanations for the splitting of the signal. ChatGPT provided numerical information for the spectra. We wanted to compare the information provided for IR, MS, and Hydrogen-1 Nuclear Magnetic Resonance (^{1}HNMR) spectra in the selected compounds to database information. The Spectral Database for Organic Compounds (SDBS) (AIST) was employed as a dependable source of data especially for MS, as our team did not have access to an MS instrument. Additionally, seven of the eleven compounds were analyzed experimentally and were selected by availability in the lab for Fourier Transformation Infrared (FTIR) (Agilent Technologies Cary 630 FTIR) and HNMR (Bruker 400 MHz Advance Core HNMR Spectrometer) spectroscopic analyses to compare the information provided by ChatGPT. The comparison of the outcomes of these experiments facilitated a comprehensive evaluation of the AI model's capabilities and accuracy in spectroscopy.

When comparing the experimental IR data collected with the major peaks and peak shapes provided by ChatGPT, there are minimal differences. While the experimental data has a set placement for the peaks, ChatGPT provides a range of values where that peak could be found. The absorption and the intensity of the peak in the experimental data was always within the range provided by ChatGPT. This observation is represented in Figure 1 for benzoic acid.



C

Summary of Expected Key Peaks:

- **O-H Stretch**: Broad, strong, 2500–3300 cm^{-1} (typically around 2800–3200 cm^{-1} for carboxylic acids).
- **C=O Stretch**: Sharp, strong, ~1700 cm^{-1} (usually around 1680–1720 cm^{-1} for carboxylic acids).
- **C-O Stretch**: 1200–1300 cm^{-1}.
- **Aromatic C-H Stretch**: 3050–3100 cm^{-1}.
- **Aromatic C=C Stretch**: 1450–1600 cm^{-1}.
- **Aromatic C-H Bending (out-of-plane)**: 690–900 cm^{-1}, with a strong band near **700 cm^{-1}** for the monosubstituted ring.

Figure 1. (A) Chemical structure of Benzoic Acid. (B) Experimental spectra of Benzoic Acid. (C) Summary: Textual output from ChatGPT-4 for prominent peaks of Benzoic Acid spectra. See Appendix for full textual output from ChatGPT for Benzoic Acid and other compounds.

Developing the Organic Chemistry II workbook

The Organic Chemistry II workbook will be composed with the specific content for Organic Chemistry II. This workbook is constructed with the input of Peer Leaders and guided by professors. The purpose of this workbook is to help practitioners have access to open-ended problems different from the problems available at the end of each chapter in a textbook.

Providing a structured workbook facilitates a structured workshop conducted by Peer Leaders. Spectroscopy is not usually covered until the second semester of Organic Chemistry at The University of Texas at El Paso and Northeastern Illinois University, although some institutions may include spectroscopy in the first semester. Organic Chemistry is notorious for being a complex topic and with the proper facilitation for Peer Leaders, the making and use of a workbook allows the best opportunities for students to reach the highest possible understanding of the subject. This has been seen with the existence of several workbooks and in the use of the Organic Chemistry I workbook (Fioramonti, et al., 2018). Our interest is to extend this to students in Organic Chemistry II by creating a workbook for the second course. We are going to utilize ChatGPT for spectral information in the development of practice problems for the Organic Chemistry II workbook.

Conclusion

ChatGPT provides reliable numerical data for MS, NMR and IR spectra of organic compounds. After successfully completing the proof of concept for ChatGPT for numerous compounds by comparing experimental and database data with the information provided by

ChatGPT, it will be used as a source of spectral information for the Organic Chemistry II workbook currently in development for PLTLIS.

ChatGPT is a resource that provides numerical spectra data that can then be used for writing the Organic Chemistry II workbook without copyright infringement. The Generative AI provides details such as the expected range, shape, and intensity of the important peaks in the desired spectra in textual output. This data could then be used to create a variety of practice problems.

After establishing the reliability of the outputs provided by ChatGPT for prominent spectra peaks, we propose that this AI tool could prove valuable in many applications, including for students. The ability of ChatGPT to provide accurate textual information about various spectra could allow students to use the generative AI as a tool to facilitate spectra interpretation. Furthermore, having established a proof of concept to determine the reliability of ChatGPT, the provided spectral information can be used across all the different topics in organic chemistry.

Acknowledgments

The authors would like to thank Annalise Gonzales, Elizabeth Lozoya, and Grecia Perea for their help in testing ChatGPT at the beginning of the project. The authors used the artificial intelligence model created by OpenAI called ChatGPT to test its reliability for providing spectra information.

References

Castro Nascimento, C. M., & Pimentel, A. S. (2023). Do Large Language Models Understand Chemistry? A Conversation with ChatGPT. *Journal of Chemical Information and Modeling*, *63*(6), 1649–1655. https://doi.org/10.1021/acs.jcim.3c00285

Christiansen, M. A. (2024, April 18). *CHATGPT assistance in creating chemistry practice problems: Pitfalls, positives, and possibilities*. ChatGPT Assistance in Creating Chemistry Practice Problems: Pitfalls, Positives, and Possibilities. https://uen.pressbooks.pub/teachingandgenerativeai/chapter/chatgpt-assistance-in-creating-chemistry-practice-problems-pitfalls-positives-and-possibilities/

Emenike, M. E., & Emenike, B. U. (2023). Was This Title Generated by ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators. *Journal of Chemical Education*, *100*(4), 1413–1418. https://doi.org/10.1021/acs.jchemed.3c00063

Fioramonti, E., Mehta, P., Gokee, A., Dooley, B., & Fraiman, A. (2018). *Organic Chemistry 1,* 2nd Edition. PLTLIS Press. https://shop.pltlis.org/

Hörnemalm, A. (2023). *ChatGPT as a software development tool: The Future of*

Development (dissertation).https://www.diva-portal.org/smash/get/diva2:1768339/FULLTEXT01.pdf

Humphry, T., & Fuller, A. L. (2023). Potential ChatGPT Use in Undergraduate Chemistry Laboratories. *Journal of Chemical Education*, *100*(4), 1434–1436. https://doi.org/10.1021/acs.jchemed.3c00006

Leon, A. J., & Vidhani, D. (2023). ChatGPT Needs a Chemistry Tutor Too. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.3c00288

Malinka, K., Peresíni, M., Firc, A., Hujnák, O., & Janus, F. (2023). On the Educational Impact of ChatGPT: Is Artificial Intelligence Ready to Obtain a University Degree? *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, *1*, 47–53. https://doi.org/10.1145/3587102.3588827

National Institute of Advanced Industrial Science and Technology (AIST). (n.d.). *Spectral Database for Organic Compounds*. SDBS, spectral database for organic compounds, AIST. https://sdbs.db.aist.go.jp/index.html

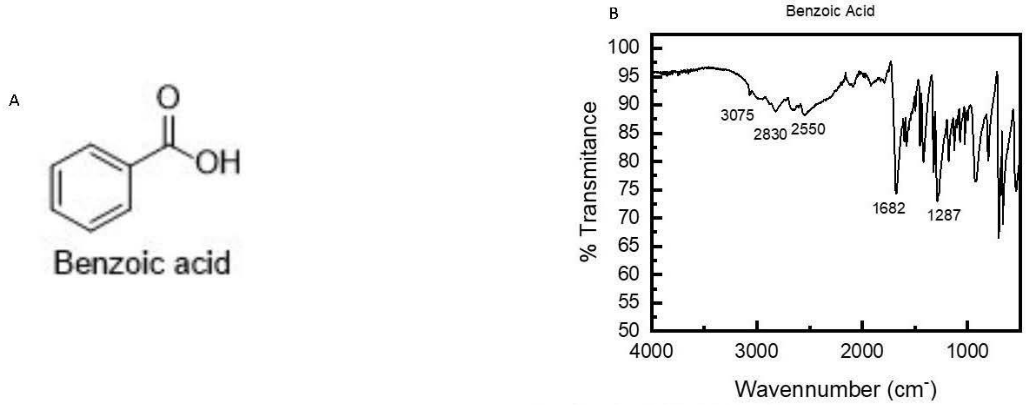
OpenAI. (2022). ChatGPT. https://chat.openai.com/chat

Rawas, S. (2024). ChatGPT: Empowering lifelong learning in the digital age of higher education. *Education and Information Technologies*, *29*(6), 6895–6908. https://doi.org/10.1007/s10639-023-12114-8

Terry, O. K. (2023, May 12). I'm a Student. You Have No Idea How Much We're Using ChatGPT. *The Chronicle of Higher Education*. https://www.chronicle.com/article/im-a-student-you-have-no-idea-how-much-were-using-chatgpt?emailConfirmed=true&supportSignUp=true&supportForgotPassword=true&email=sdelga7701%40gmail.com&success=true&code=success&bc_nonce=aig7yhsb8mcgts137g9655

Appendix

Figures S1-S33 - Full Output from ChatGPT of Eleven Compounds



C

Certainly! Here's a detailed description of the expected IR (infrared) spectra for **benzoic acid** based on its functional groups and molecular structure.

Structure of Benzoic Acid (C_6H_5COOH):

Benzoic acid consists of a benzene ring attached to a carboxylic acid group (-COOH). The main functional groups contributing to the IR spectrum include:

- The **aromatic ring** (C=C bonds in the benzene ring).
- The **carboxylic acid group** (-COOH), which includes:
  - A carbonyl (C=O) bond.
  - A hydroxyl (O-H) group.

Key Peaks in the IR Spectrum of Benzoic Acid:

1. **O-H Stretch (Broad, 2500–3300 cm^{-1}):**
   - The **hydroxyl group** (-OH) in the carboxylic acid produces a very broad and strong absorption band.
   - This broadening happens due to hydrogen bonding between carboxylic acid molecules, forming dimers in solid and liquid phases.
   - The range for carboxylic acid O-H stretching is broader and more intense than for alcohols or phenols.
2. **C=O Stretch (Sharp, Strong, ~1700 cm^{-1}):**
   - The **carbonyl (C=O) stretch** of the carboxylic acid typically appears around **1700 cm^{-1}**.
   - In benzoic acid, this peak is generally slightly **lowered** (around 1680–1720 cm^{-1}) compared to a simple ketone due to resonance and the electron-withdrawing effects of the benzene ring.
3. **C-O Stretch (1200–1300 cm^{-1}):**
   - This stretch represents the **C-O bond** in the carboxylic acid group and appears as a strong peak in the **1200–1300 cm^{-1}** region.
   - It is often paired with the O-H bending vibration.
4. **Aromatic C-H Stretch (3050–3100 cm^{-1}):**
   - The **aromatic C-H bonds** in the benzene ring produce a weak to moderate peak in the **3050–3100 cm^{-1}** region.
   - These are sharper and less intense compared to the O-H stretch.
5. **C=C Stretch in the Aromatic Ring (1450–1600 cm^{-1}):**
   - The stretching vibrations of the **C=C bonds** in the aromatic ring generally produce several weak to moderate peaks in the **1450–1600 cm^{-1}** range.
   - These are characteristic of the conjugated double bonds in the benzene ring and are often split into two or more peaks due to the symmetry of the ring.
6. **Aromatic C-H Bending (690–900 cm^{-1}):**
   - **Out-of-plane bending** vibrations of the aromatic C-H bonds typically appear as strong peaks between **690–900 cm^{-1}**.
   - For benzoic acid, expect a significant peak near **700 cm^{-1}**, characteristic of the monosubstituted benzene ring.

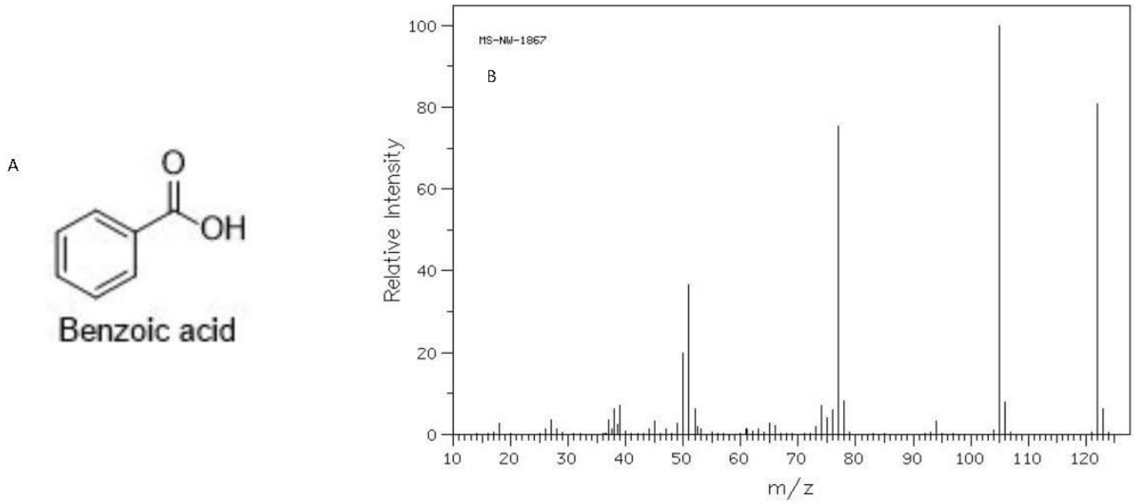
Summary of Expected Key Peaks:

- **O-H Stretch**: Broad, strong, 2500–3300 cm^{-1} (typically around 2800–3200 cm^{-1} for carboxylic acids).
- **C=O Stretch**: Sharp, strong, ~1700 cm^{-1} (usually around 1680–1720 cm^{-1} for carboxylic acids).
- **C-O Stretch**: 1200–1300 cm^{-1}.
- **Aromatic C-H Stretch**: 3050–3100 cm^{-1}.
- **Aromatic C=C Stretch**: 1450–1600 cm^{-1}.
- **Aromatic C-H Bending (out-of-plane)**: 690–900 cm^{-1}, with a strong band near **700 cm^{-1}** for the monosubstituted ring.

Additional Considerations:

- **Hydrogen Bonding**: In solid or concentrated solutions, benzoic acid forms dimers via hydrogen bonding. This interaction causes the broadening of the O-H stretch.
- **Resonance**: The conjugation between the carboxyl group and the benzene ring slightly alters the position of the C=O stretching vibration compared to non-conjugated carboxylic acids.

Figure S1: (A) Chemical structure of Benzoic Acid. (B) Experimental spectra of Benzoic Acid. (C) Textual output from ChatGPT-4 for prominent peaks of Benzoic Acid spectra.



C

Memory updated

Sure! Here's a more detailed explanation of the fragmentation process of benzoic acid in mass spectrometry:

1. Molecular Ion (M^+):

- The molecular ion is formed when the benzoic acid molecule ($C_7H_6O_2$) loses an electron, resulting in a positively charged ion.
- **m/z = 122** corresponds to the molecular ion, which represents the entire benzoic acid molecule.

2. Loss of Carboxyl Group (COOH):

- One of the most common fragmentation patterns for benzoic acid is the loss of the carboxy group (-COOH).
- This leaves behind a phenyl ion ($C_6H_5^+$), which is very stable due to the aromatic structure.
- **m/z = 105** corresponds to the phenyl ion ($C_6H_5^+$), which is a significant fragment in the spectrum.

Reaction:

$$C_6H_5COOH^+ \rightarrow C_6H_5^+ + COOH$$

3. Cleavage within the Benzene Ring:

- Further fragmentation of the phenyl ion can occur, leading to the loss of small groups (like -CH_3 or -C_2H_2), though this is less common.
- One of the resulting fragments is **C_6H_5 → $C_6H_5^+$**, leading to peaks such as **m/z = 77** ($C_6H_5^+$). This corresponds to a stable aromatic phenyl ring fragment.

4. Smaller Fragments:

- Smaller fragments such as **m/z = 51** can be produced, which might correspond to further cleavage of the benzene ring. This can occur when bonds within the ring break, leading to smaller hydrocarbon chains.

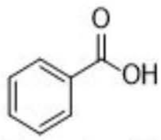
Summary of Key Peaks:

- **m/z = 122**: Molecular ion peak (M^+, $C_7H_6O_2$).
- **m/z = 105**: Loss of COOH ($C_6H_5^+$).
- **m/z = 77**: Phenyl ion ($C_6H_5^+$).
- **m/z = 51**: Smaller fragments from ring cleavage.

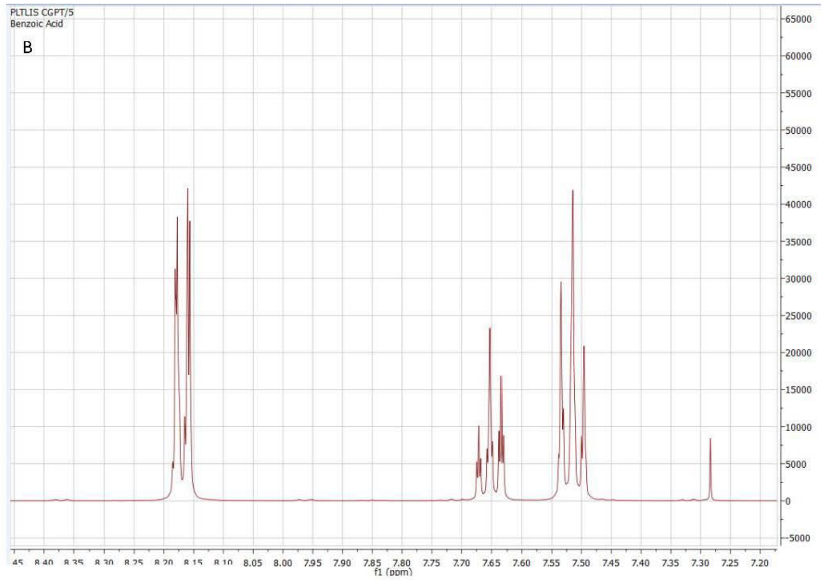
This stepwise breakdown provides insight into the major fragments that form when benzoic acid is ionized and fragmented during a mass spectrometry experiment.

Figure S2: (A) Chemical structure of Benzoic Acid. (B) SDBS Database of Benzoic Acid (11). (C) Textual output from ChatGPT-4 for prominent fragments of Benzoic Acid spectra.

A



Benzoic acid



C

The ^1H-NMR spectrum for benzoic acid (C_6H_5COOH) can be analyzed using the n+1 rule, where **n** is the number of neighboring protons and **n+1** gives the number of peaks (multiplets) for each signal. Here's how the protons in benzoic acid split according to the n+1 rule:

Benzoic Acid Structure:

- The structure consists of a benzene ring with a carboxylic acid group ($-COOH$) attached to it.
- The carboxylic acid proton does not split because it is not adjacent to any other protons.

Proton Environment Breakdown:

1. **Carboxylic Acid Proton** ($-COOH$):
   - This proton usually appears as a singlet because there are no protons directly adjacent to it for coupling.
   - Chemical shift: Typically appears around 10-13 ppm.
2. **Aromatic Protons on the Benzene Ring:**
   - The benzene ring protons are divided into three distinct environments due to the presence of the electron-withdrawing $-COOH$ group, which affects their chemical shift:
     - **Ortho Protons (2 protons):** These protons are adjacent to the carboxylic acid group, meaning they will experience some splitting from the other aromatic protons.
     - **Meta Protons (2 protons):** These are two carbons away from the carboxylic group.
     - **Para Proton (1 proton):** This is directly opposite the carboxylic acid group.

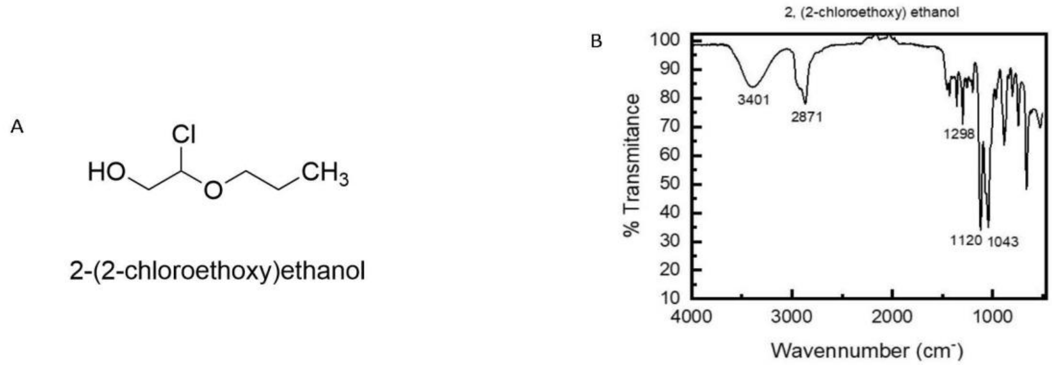
Approximate Splitting Pattern:

- **Ortho Protons (H at positions 2 and 6):**
  - These protons will have three neighbors (meta and para protons), leading to a splitting pattern of 3+1 = 4 peaks (quartet).
  - Chemical shift: Around 7.5-8 ppm.
- **Meta Protons (H at positions 3 and 5):**
  - These protons have two ortho neighbors, so they will exhibit 2+1 = 3 peaks (triplet).
  - Chemical shift: Around 7.3-7.5 ppm.
- **Para Proton (H at position 4):**
  - This proton has two meta neighbors, so it will exhibit 2+1 = 3 peaks (triplet).
  - Chemical shift: Around 7.3 ppm.

Summary of Splitting Patterns and Chemical Shifts:

- **Carboxylic Acid Proton:** Singlet, ~10-13 ppm.
- **Ortho Protons:** Quartet, ~7.5-8 ppm.
- **Meta Protons:** Triplet, ~7.3-7.5 ppm.
- **Para Proton:** Triplet, ~7.3 ppm.

Figure S3: (A) Chemical structure of Benzoic Acid. (B) Experimental spectra of Benzoic Acid. (C) Textual output from ChatGPT-4 for prominent peaks of Benzoic Acid spectra.



C

Certainly! Let's break down the IR spectrum for **2-(2-Chloroethoxy)ethanol ($C_4H_9ClO_2$)**. This compound contains two functional groups: an **ether (-O-)** and a **hydroxyl (-OH)** group, along with a **chloroalkane (C-Cl)** bond.

Structure of 2-(2-Chloroethoxy)ethanol:

- **Ether group (-O-)**: An oxygen atom linking two ethyl chains.
- **Hydroxyl group (-OH)**: At the end of one ethyl chain.
- **Chloro group (-Cl)**: Attached to the other ethyl chain.

Key Peaks Expected in the IR Spectrum:

1. **O-H Stretch (Broad, 3200–3600 cm^{-1}):**
   - The hydroxyl (-OH) group leads to a **broad, strong absorption band** in the **3200–3600 cm^{-1}** region.
   - The broadness is due to hydrogen bonding, particularly if the sample is in the liquid state or highly concentrated.
2. **C-H Stretch (Alkyl, 2850–3000 cm^{-1}):**
   - The C-H bonds in the **methylene (-CH_2)** groups from both the ethanol and chloroethoxy chains give rise to **moderate** peaks in the **2850–3000 cm^{-1}** range.
   - These are typical of alkyl C-H stretching vibrations, with symmetric and asymmetric stretches.
3. **C-O Stretch (Ether, 1050–1150 cm^{-1}):**
   - The **C-O** bond in the ether group will show a peak in the **1050–1150 cm^{-1}** region.
   - This is characteristic of an ether and typically appears as a strong peak due to the oxygen's impact on bond polarity.
4. **C-O Stretch (Alcohol, 1000–1100 cm^{-1}):**
   - The hydroxyl group's **C-O bond** also has a characteristic peak in the **1000–1100 cm^{-1}** region, sometimes overlapping with the ether C-O stretch.
   - This peak can vary slightly based on hydrogen bonding and solvent interactions.
5. **C-Cl Stretch (600–800 cm^{-1}):**
   - The **C-Cl bond** leads to a peak in the **600–800 cm^{-1}** region, often around **700 cm^{-1}**.
   - This absorption is typically moderate in intensity and sharp.
6. **C-H Bending (1350–1500 cm^{-1}):**
   - The **bending vibrations** of the alkyl C-H bonds produce moderate peaks in the **1350–1500 cm^{-1}** region.
   - These are common for methylene groups in aliphatic chains.
7. **O-H Bending (1400–1450 cm^{-1}):**
   - A weaker peak due to **O-H bending** may be observed in the **1400–1450 cm^{-1}** range.

Summary of Key Peaks:

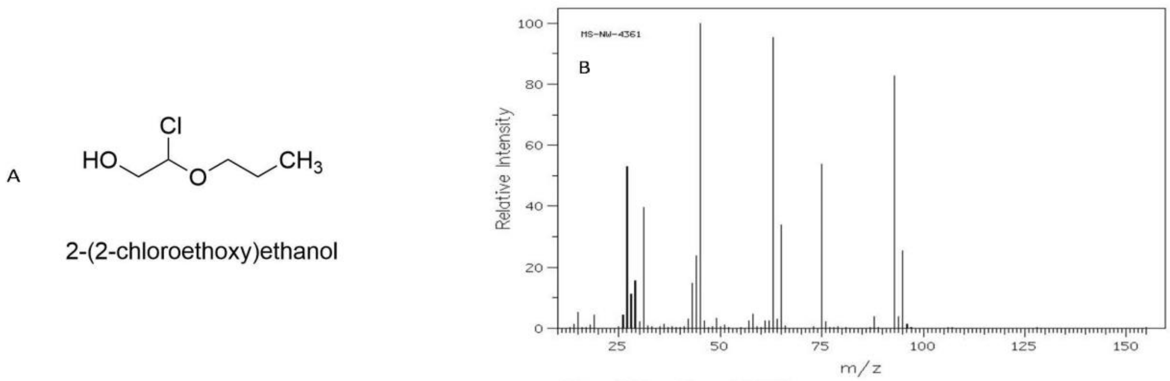
- **O-H Stretch (Hydroxyl)**: Broad, strong, 3200–3600 cm^{-1}.
- **C-H Stretch (Alkyl)**: 2850–3000 cm^{-1}.
- **C-O Stretch (Ether)**: 1050–1150 cm^{-1}.
- **C-O Stretch (Alcohol)**: 1000–1100 cm^{-1}.
- **C-Cl Stretch (Chloroalkane)**: 600–800 cm^{-1}.
- **C-H Bending**: 1350–1500 cm^{-1}.
- **O-H Bending**: 1400–1450 cm^{-1} (if detectable).

Additional Considerations:

- **Hydrogen bonding**: The hydroxyl group can form hydrogen bonds, which may broaden and shift the O-H stretching peak.
- **Chlorine's electronegativity**: The presence of chlorine slightly impacts nearby vibrations, particularly the C-H stretches, but these effects are generally subtle.

This IR profile gives a clear representation of the different functional groups present in 2-(2-Chloroethoxy)ethanol. If you need more details or clarification, feel free to ask!

Figure S4: (A) Chemical structure of 2-(2-Chloroethoxy) ethanol. (B) Experimental spectra of 2-(2-Chloroethoxy) ethanol. (C) Textual output from ChatGPT-4 for prominent peaks of 2-(2-Chloroethoxy) ethanol spectra.



C

Mass Spectrometry Analysis for 2-(2-Chloroethoxy)ethanol ($C_4H_9ClO_2$)

Molecular Formula: $C_4H_9ClO_2$

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 4 = 48.04
- Hydrogen (H): 1.008 × 9 = 9.072
- Chlorine (Cl): 35.45 × 1 = 35.45
- Oxygen (O): 16.00 × 2 = 32.00

Total = 124.56 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M⁺):

- The molecular ion represents the intact molecule, where an electron is removed. The molecular ion peak would appear at **m/z = 124**.
- However, since the molecule contains chlorine, two isotopes of chlorine (^35Cl and ^37Cl) are present. Chlorine causes a characteristic isotopic pattern with two peaks:
  - **m/z = 124** for the ^35Cl isotope (about 75% abundance).
  - **m/z = 126** for the ^37Cl isotope (about 25% abundance).

This pattern helps confirm the presence of chlorine in the molecule.

2. Loss of Hydroxyl Group (-OH):

- One possible fragmentation pathway is the loss of a hydroxyl group (-OH), which would result in a fragment with a mass loss of 17 Da.
- This gives a fragment at **m/z = 107**.

  Reaction:

$$C_4H_9ClO_2^+ \rightarrow C_4H_8ClO^+ + OH$$

3. Loss of Ethanol Group (-C_2H_5O):

- Another common fragmentation involves the loss of the ethanol portion (-C_2H_5O, 45 Da), leaving behind a fragment at **m/z = 79**. This corresponds to the remaining portion of the molecule with the chlorine atom still attached.

  Reaction:

$$C_4H_9ClO_2^+ \rightarrow C_2H_4Cl^+ + C_2H_5O$$

4. Cleavage of the Chloroethoxy Chain:

- The molecule can also fragment at the bond between the carbon and the oxygen of the chloroethoxy group (-CH_2CH_2Cl). This cleavage would result in two major fragments:
  - **m/z = 93**: Loss of a -CH_2CH_2Cl group (31 Da), giving the fragment $C_2H_5O_2$^+.
  - **m/z = 31**: A smaller fragment corresponding to the remaining CH_2OH (methyl hydroxyl) group.

  Reaction:

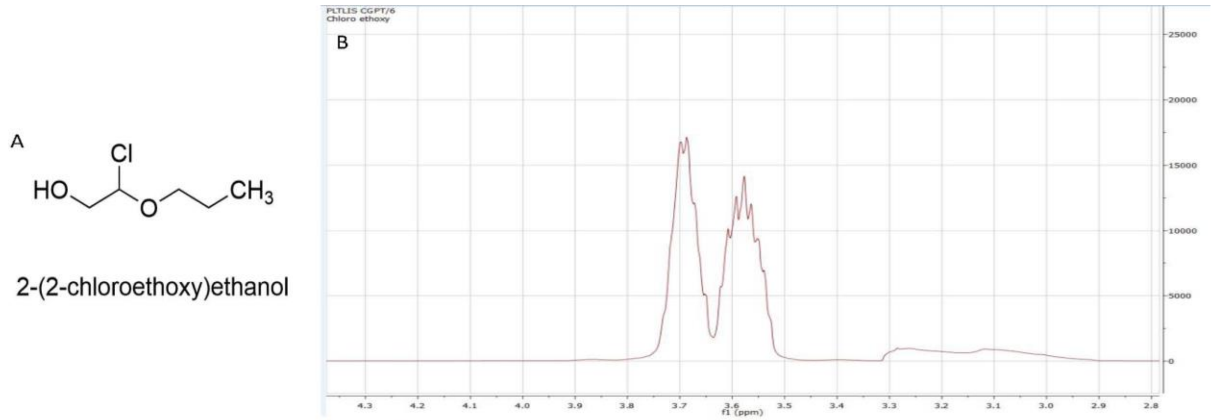
$$C_4H_9ClO_2^+ \rightarrow C_2H_5O_2^+ + CH_2CH_2Cl$$

Summary of Expected Peaks:

- **m/z = 124 (M⁺)**: Molecular ion (with isotopic peaks at **m/z = 126** due to chlorine isotopes).
- **m/z = 107**: Loss of the hydroxyl group (-OH).
- **m/z = 93**: Cleavage of the chloroethoxy group (-CH_2CH_2Cl).
- **m/z = 79**: Loss of the ethanol group (-C_2H_5O).
- **m/z = 31**: A smaller fragment corresponding to the -CH_2OH group.

These fragmentation patterns give a clear indication of the structure of **2-(2-Chloroethoxy)ethanol**, allowing you to identify key fragments in the mass spectrum.

Figure S5: (A) Chemical structure of 2-(2-Chloroethoxy) ethanol. (B) SDBS Database of 2-(2-Chloroethoxy) ethanol (11) (C) Textual output from ChatGPT-4 for prominent fragments of 2-(2-Chloroethoxy) ethanol spectra.



C

For the compound **2-(2-Chloroethoxy) ethanol** ($ClCH_2CH_2OCH_2CH_2OH$), we can use the **n+1 rule** to determine the splitting pattern in the ^1H-NMR spectrum.

Structure:

The structure consists of two main groups:

1. A **chloroethyl** group ($ClCH_2CH_2-$)
2. A **hydroxyethyl** group ($HOCH_2CH_2-$) These groups are connected by an oxygen atom, forming an ether linkage between the two.

Proton Environments:

The protons are in several distinct environments:

1. **Hydroxyl proton ($-OH$):**
   - This proton typically appears as a broad singlet because it does not couple with neighboring protons (due to rapid exchange with solvent).
   - Chemical shift: Around 2-4 ppm (varies depending on solvent and concentration).
2. **Protons on the -CH_2 group adjacent to the hydroxyl group ($HOCH_2-$):**
   - This methylene group is directly connected to the hydroxyl group and to another methylene group.
   - It will be split by the protons on the adjacent -CH_2O- group (2 neighboring protons), giving a 2+1 = 3 peak (triplet).
   - Chemical shift: Around 3.4-4 ppm (due to proximity to oxygen).
3. **Protons on the -CH_2O- group (next to ether oxygen):**
   - This methylene group is adjacent to both the oxygen and another -CH_2Cl group.
   - It will be split by the protons on the adjacent -CH_2OH group and the -CH_2Cl group. Since each adjacent methylene has 2 protons, it will experience a splitting of (2+2)+1 = 5 peaks (quintet).
   - Chemical shift: Around 3.6-4 ppm.
4. **Protons on the -CH_2Cl group:**
   - This methylene group is adjacent to the -CH_2O- group.
   - It will experience splitting from the two protons on the adjacent -CH_2O- group, giving a 2+1 = 3 peak (triplet).
   - Chemical shift: Around 3.5-3.7 ppm (due to the influence of the electronegative chlorine atom).
5. **Protons on the -CH_2- group adjacent to the chlorine atom:**
   - This methylene group will be split by the two protons on the adjacent -CH_2Cl group, giving a 2+1 = 3 peak (triplet).
   - Chemical shift: Around 3.3-3.5 ppm.

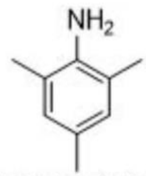
Summary of Proton Splitting and Chemical Shifts:

- **Hydroxyl Proton ($-OH$)**: Singlet (broad), ~2-4 ppm.
- $HOCH_2-$ **Proton**: Triplet, ~3.4-4 ppm.
- -CH_2O- **Proton**: Quintet, ~3.6-4 ppm.
- -CH_2Cl Proton: Triplet, ~3.5-3.7 ppm.
- -CH_2- Proton next to chlorine: Triplet, ~3.3-3.5 ppm.

This gives you a general overview of how the protons in 2-(2-chloroethoxy) ethanol would behave in an ^1H-NMR spectrum using the n+1 rule.

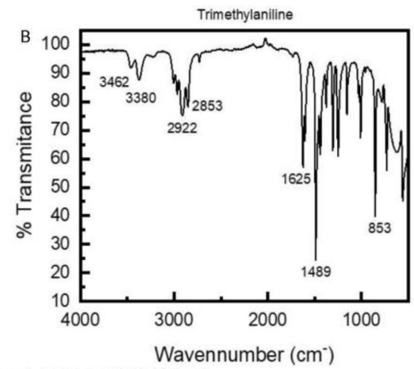
Figure S6: (A) Chemical structure of 2-(2-Chloroethoxy) ethanol. (B) Experimental spectra of 2-(2-Chloroethoxy) ethanol. (C) Textual output from ChatGPT-4 for prominent peaks of 2-(2-Chloroethoxy) ethanol spectra.

A



2,4,6-trimethylanaline

B



C

Certainly! Let's break down the expected IR spectrum for **2,4,6-Trimethylaniline ($C_9H_{13}N$)**, which is an aromatic amine with three methyl groups attached to the benzene ring.

Structure of 2,4,6-Trimethylaniline:

- **Aniline group**: A benzene ring with an amino group (-NH_2) attached at the **1-position**.
- **Methyl groups**: Three methyl (-CH_3) groups attached at the **2-, 4-, and 6-positions** on the benzene ring.

Key Functional Groups:

1. **Amino group (-NH_2)**.
2. **Aromatic ring (benzene)**.
3. **Methyl (-CH_3) groups** attached to the ring.

Key Peaks Expected in the IR Spectrum of 2,4,6-Trimethylaniline:

1. **N-H Stretch (Amino, 3300–3500 cm^{-1}):**
   - The **primary amine (-NH_2)** produces two peaks in the **3300–3500 cm^{-1}** range due to **symmetric and asymmetric N-H stretching**.
   - These peaks are generally **moderate to weak** in intensity and somewhat sharp.
   - In 2,4,6-Trimethylaniline, the amino group may exhibit some slight shifts in frequency due to the influence of the methyl groups.
2. **C-H Stretch (Aromatic, 3000–3100 cm^{-1}):**
   - The **C-H bonds** in the aromatic ring will show **weak to moderate peaks** in the **3000–3100 cm^{-1}** range.
   - These peaks arise from the stretching of the aromatic C-H bonds in the benzene ring.
3. **C-H Stretch (Methyl groups, 2850–2950 cm^{-1}):**
   - The **C-H bonds** in the **methyl (-CH_3) groups** attached to the benzene ring produce **moderate peaks** in the **2850–2950 cm^{-1}** region.
   - These are typical alkyl C-H stretching vibrations and will be stronger than the aromatic C-H stretches.
4. **N-H Bending (Amino, 1600–1650 cm^{-1}):**
   - A strong **N-H bending** vibration (scissoring) occurs in the **1600–1650 cm^{-1}** range.
   - This peak is typical of primary amines and is often **moderate to strong** in intensity.
5. **Aromatic C=C Stretch (1450–1600 cm^{-1}):**
   - The **C=C stretching** vibrations of the benzene ring appear as a series of weak to moderate peaks in the **1450–1600 cm^{-1}** range.
   - These peaks correspond to the symmetric and asymmetric stretches of the conjugated double bonds in the aromatic ring.
6. **C-H Bending (Methyl, 1370–1470 cm^{-1}):**
   - The **bending vibrations** of the methyl groups, particularly **CH_3 bending (deformation)**, produce peaks around **1370 cm^{-1}** and **1460 cm^{-1}**.
   - These peaks are typically **strong** due to the three methyl groups.
7. **Aromatic C-H Out-of-Plane Bending (700–900 cm^{-1}):**
   - The **out-of-plane bending** of the aromatic C-H bonds produces characteristic peaks in the **700–900 cm^{-1}** region.
   - For 2,4,6-trisubstituted benzene derivatives, there are typically two distinct peaks in this region (often around **750–850 cm^{-1}**), indicative of the substitution pattern.

Summary of Key Peaks:

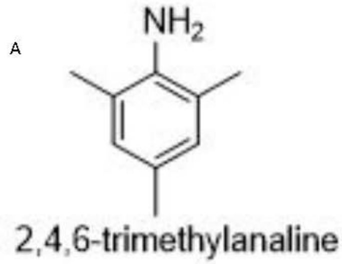
- **N-H Stretch (Amino)**: 3300–3500 cm^{-1} (two peaks, moderate intensity).
- **C-H Stretch (Aromatic)**: 3000–3100 cm^{-1} (weak to moderate).
- **C-H Stretch (Methyl)**: 2850–2950 cm^{-1} (moderate intensity).
- **N-H Bending (Amino)**: 1600–1650 cm^{-1} (strong).
- **Aromatic C=C Stretch**: 1450–1600 cm^{-1} (weak to moderate).
- **C-H Bending (Methyl)**: 1370–1470 cm^{-1} (strong).
- **Aromatic C-H Out-of-Plane Bending**: 700–900 cm^{-1} (moderate, characteristic for trisubstituted benzenes).

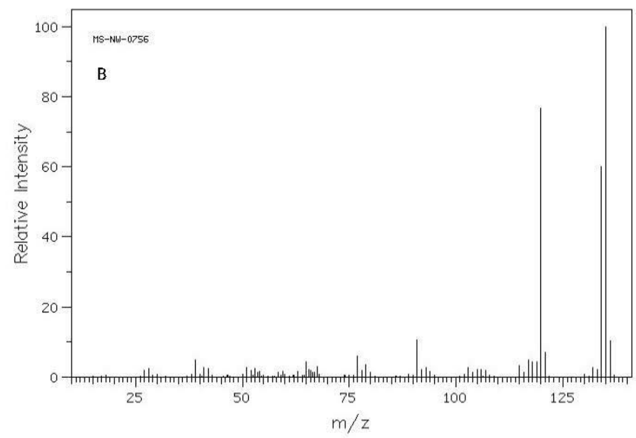
Additional Considerations:

- **Electron donation**: The methyl groups attached to the benzene ring are electron-donating, which slightly shifts the stretching frequencies of the aromatic C=C bonds compared to an unsubstituted aniline.
- **Substitution pattern**: The 2,4,6-trimethyl substitution pattern on the benzene ring will also influence the position and intensity of the out-of-plane bending peaks, giving a unique signature for this compound.

This analysis should give you a detailed understanding of the IR spectrum for 2,4,6-Trimethylaniline. Let me know if you'd like further clarification!

Figure S7: (A) Chemical structure of 2,4,6-Trimethylaniline. (B) Experimental spectra of 2,4,6-Trimethylaniline. (C) Textual output from ChatGPT-4 for prominent peaks of 2,4,6-Trimethylaniline spectra.





C

Mass Spectrometry Analysis for 2,4,6-Trimethylaniline ($C_9H_{13}N$)

Molecular Formula: $C_9H_{13}N$
Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 9 = 108.09
- Hydrogen (H): 1.008 × 13 = 13.104
- Nitrogen (N): 14.01 × 1 = 14.01

Total = 135.20 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M⁺):

- The molecular ion peak, which corresponds to the intact molecule, appears at **m/z = 135**. This peak indicates the complete molecule.

2. Loss of Methyl Groups (-CH_3):

- 2,4,6-Trimethylaniline has three methyl (-CH_3) groups attached to the aromatic ring, and one of the common fragmentation pathways is the sequential loss of these groups.
- Loss of one methyl group results in a fragment at **m/z = 120**.
- Loss of a second methyl group results in a fragment at **m/z = 105**.
- Loss of the third methyl group results in a fragment at **m/z = 90**.

Reactions:

$$C_9H_{13}N^+ \rightarrow C_8H_{10}N^+ + CH_3 \quad (m/z = 120)$$
$$C_8H_{10}N^+ \rightarrow C_7H_7N^+ + CH_3 \quad (m/z = 105)$$
$$C_7H_7N^+ \rightarrow C_6H_4N^+ + CH_3 \quad (m/z = 90)$$

3. Cleavage of the Amino Group (-NH_2):

- The molecule can also undergo fragmentation with the loss of the amino group (-NH_2), leading to a fragment at **m/z = 118**.

Reaction:

$$C_9H_{13}N^+ \rightarrow C_9H_{12}^+ + NH_2 \quad (m/z = 118)$$

4. Cleavage of the Methyl and Amino Groups Together:

- Another pathway involves the loss of both the methyl and amino groups, resulting in a fragment at **m/z = 103**.

Reaction:

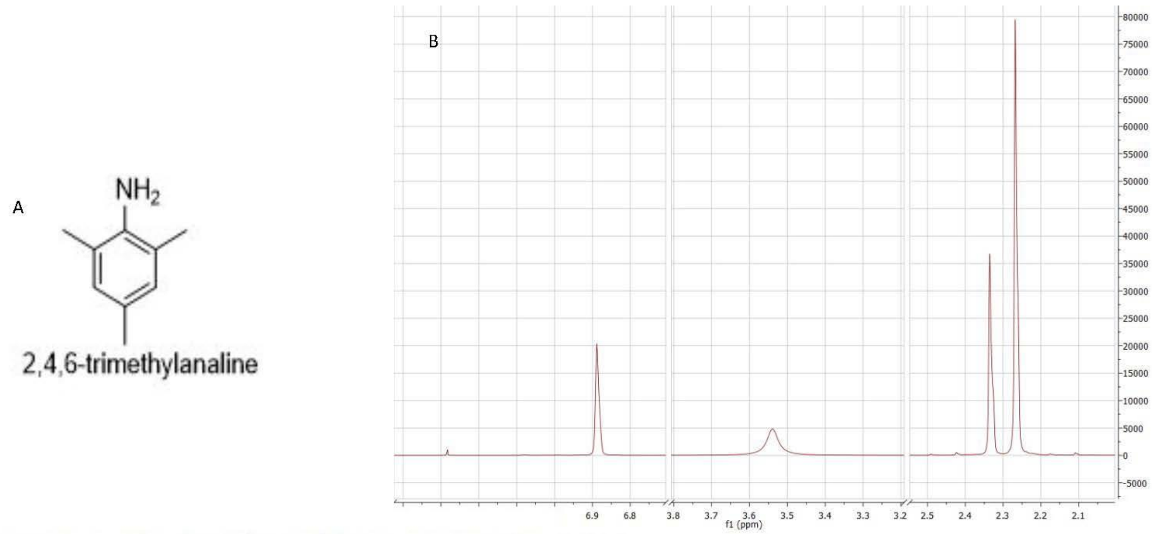
$$C_9H_{13}N^+ \rightarrow C_7H_9^+ + CH_3 + NH_2 \quad (m/z = 103)$$

Summary of Expected Peaks:

- **m/z = 135 (M⁺)**: Molecular ion peak.
- **m/z = 120**: Loss of one methyl group (-CH_3).
- **m/z = 105**: Loss of two methyl groups.
- **m/z = 90**: Loss of all three methyl groups.
- **m/z = 118**: Loss of the amino group (-NH_2).
- **m/z = 103**: Loss of both the amino and methyl groups.

These fragmentation patterns help identify the structure of **2,4,6-Trimethylaniline** in a mass spectrum by highlighting the major pathways and fragments.

Figure S8: (A) Chemical structure of 2,4,6-Trimethylaniline. (B) SDBS Database of 2,4,6-Trimethylaniline (11) (C) Textual output from ChatGPT-4 for prominent fragments of 2,4,6-Trimethylaniline spectra.



C

For **2,4,6-Trimethylaniline** (chemical formula: $C_9H_{13}N$), the ^1H-NMR spectrum can be analyzed using the **n+1 rule**. Here's a breakdown of the proton environments and their splitting patterns.

Structure of 2,4,6-Trimethylaniline:

- It consists of an aniline group (-NH_2) attached to a benzene ring, with **three methyl groups** at the **2, 4, and 6** positions of the ring (relative to the amine group).

Proton Environments:

1. **Amino group protons** (-NH_2):
   - The protons of the amino group do not split because they are not adjacent to any protons that can couple with them.
   - This will appear as a broad singlet.
   - Chemical shift: Around 3.5-4.5 ppm (depends on hydrogen bonding and solvent).
2. **Aromatic protons**:
   - There are **two aromatic protons** on the benzene ring (at positions 3 and 5) because positions 2, 4, and 6 are occupied by methyl groups, which makes the ring symmetric.
   - These two protons are in the **same chemical environment** due to the symmetry of the molecule.
   - These protons will split each other, following the n+1 rule. Each proton has one neighbor (the other aromatic proton), resulting in a **doublet** (1+1 = 2 peaks).
   - Chemical shift: Around 6.5-7.5 ppm (typical of aromatic protons).
3. **Methyl group protons** (-CH_3):
   - There are **three equivalent methyl groups** at positions 2, 4, and 6 on the benzene ring. These groups are not adjacent to any protons, as their neighboring carbons are part of the benzene ring.
   - Each methyl group will appear as a **singlet** (no adjacent protons to cause splitting).
   - Chemical shift: Around 2.0-2.3 ppm.

Summary of Proton Splitting and Chemical Shifts:

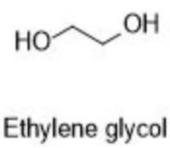
1. **Amino group protons** (-NH_2): Broad singlet, ~3.5-4.5 ppm.
2. **Aromatic protons (H at positions 3 and 5)**: Doublet, ~6.5-7.5 ppm.
3. **Methyl protons** (-CH_3 **at positions 2, 4, and 6)**: Singlet, ~2.0-2.3 ppm.

Overall Splitting Pattern:

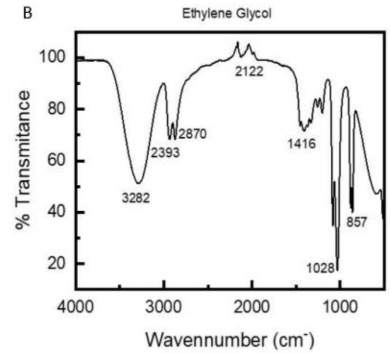
- -NH_2: Broad singlet.
- **Aromatic protons**: Doublet.
- **Methyl groups**: Singlet.

Figure S9: (A) Chemical structure of 2,4,6-Trimethylaniline. (B) Experimental spectra of 2,4,6-Trimethylaniline. (C) Textual output from ChatGPT-4 for prominent peaks of 2,4,6-Trimethylaniline spectra.

A



B



C

For **ethylene glycol ($C_2H_6O_2$)**, the IR spectrum will be dominated by the presence of **two hydroxyl (-OH)** groups and **C-H** bonds from the ethylene backbone.

Structure of Ethylene Glycol:

- **Hydroxyl groups (-OH)**: Two hydroxyl groups, one on each end of the ethylene chain.
- **Alkyl backbone ($-CH_2-CH_2-$)**: Two methylene ($-CH_2-$) groups connecting the hydroxyls.

Key Peaks Expected in the IR Spectrum of Ethylene Glycol:

1. **O-H Stretch (Broad, 3200–3600 cm^{-1})**:
   - The **hydroxyl (-OH)** groups produce a **broad, strong absorption band** in the **3200–3600 cm^{-1}** region.
   - This broadening is typical of alcohols due to **hydrogen bonding**, which is especially strong in ethylene glycol because it has two hydroxyl groups, leading to extensive intermolecular hydrogen bonding.
2. **C-H Stretch (Alkyl, 2850–3000 cm^{-1})**:
   - The **C-H bonds** from the two methylene ($-CH_2-$) groups give rise to **moderate peaks** in the **2850–3000 cm^{-1}** range.
   - These are characteristic alkyl C-H stretching vibrations, typically including both symmetric and asymmetric stretches.
3. **O-H Bending (1400–1440 cm^{-1})**:
   - There is a weaker, yet noticeable peak due to **O-H bending (scissoring)** in the **1400–1440 cm^{-1}** range.
   - This corresponds to the bending of the hydroxyl groups.
4. **C-O Stretch (1050–1150 cm^{-1})**:
   - The **C-O** stretching vibrations from the alcohol functional groups produce **strong peaks** in the **1050–1150 cm^{-1}** region.
   - This peak is typically sharp and intense due to the two hydroxyl groups, and often appears as one of the strongest features in the spectrum.
5. **C-H Bending (1350–1470 cm^{-1})**:
   - The **bending vibrations** of the methylene ($-CH_2-$) groups produce **moderate peaks** in the **1350–1470 cm^{-1}** range, typically related to C-H scissoring and deformation.
   - These peaks are commonly found for alkyl chains.
6. **O-H Out-of-Plane Bending (650–700 cm^{-1})**:
   - There may also be a **bending vibration** for the O-H bond in the out-of-plane direction in the **650–700 cm^{-1}** region, though it is usually weaker.

Summary of Key Peaks:

- **O-H Stretch (Hydroxyl)**: Broad, strong, 3200–3600 cm^{-1}.
- **C-H Stretch (Alkyl)**: 2850–3000 cm^{-1}.
- **O-H Bending**: 1400–1440 cm^{-1} (weak).
- **C-O Stretch**: 1050–1150 cm^{-1} (strong).
- **C-H Bending**: 1350–1470 cm^{-1} (moderate).
- **O-H Out-of-Plane Bending**: 650–700 cm^{-1} (if detectable).

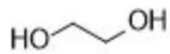
Additional Considerations:

- **Hydrogen bonding**: Ethylene glycol's extensive hydrogen bonding leads to broadening and sometimes shifting of the O-H stretching peak. This peak may appear even broader than in other alcohols due to the presence of two hydroxyl groups.
- **Symmetry**: Ethylene glycol is a relatively simple symmetrical molecule, and this symmetry affects how the peaks for similar groups (like the two hydroxyls) combine, often giving clearer, more intense peaks compared to more complex alcohols.

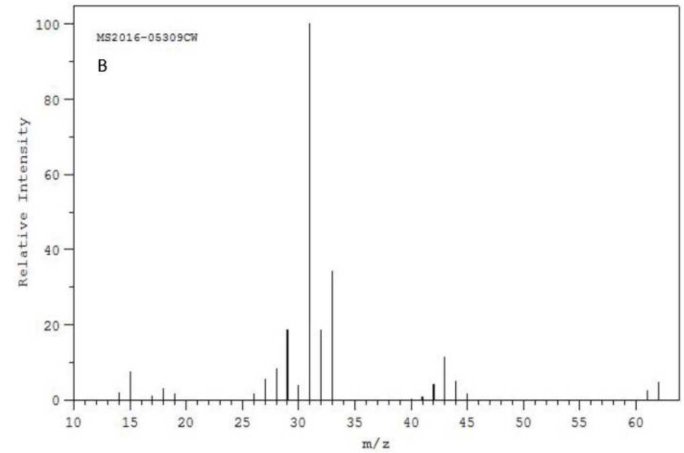
This profile provides a detailed understanding of the IR spectrum for ethylene glycol. Let me know if you'd like any further details!

Figure S10: (A) Chemical structure of Ethylene Glycol. (B) Experimental spectra of Ethylene Glycol. (C) Textual output from ChatGPT-4 for prominent peaks of Ethylene Glycol spectra

A



Ethylene glycol



C

Mass Spectrometry Analysis for Ethylene Glycol ($C_2H_6O_2$)

Molecular Formula: $C_2H_6O_2$

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 2 = 24.02
- Hydrogen (H): 1.008 × 6 = 6.048
- Oxygen (O): 16.00 × 2 = 32.00

Total = 62.07 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M⁺):

- The molecular ion peak, representing the intact molecule, appears at **m/z = 62**. However, ethylene glycol is prone to fragmentation, and the molecular ion is often weak or absent.

2. Loss of Hydroxyl Group (-OH):

- A common fragmentation pathway involves the loss of one hydroxyl group (-OH), resulting in a fragment at **m/z = 45**. This represents the loss of 17 Da (hydroxyl group).

  Reaction:

$$C_2H_6O_2^+ \rightarrow C_2H_5O^+ + OH$$

3. Cleavage of Carbon-Carbon Bond:

- The molecule can also fragment via cleavage of the C–C bond between the two carbon atoms, producing two distinct fragments:
  - **m/z = 31**: This corresponds to the hydroxymethyl cation (CH_2OH^+).
  - **m/z = 31**: A similar fragment results from the other hydroxyl group, leading to two identical fragments in this case.

  Reaction:

$$C_2H_6O_2^+ \rightarrow CH_2OH^+ + CH_2OH^+$$

4. Loss of Water (H_2O):

- Another fragmentation pathway involves the loss of a water molecule (18 Da), which gives a fragment at **m/z = 44**. This indicates the loss of both hydroxyl groups together as a water molecule.

  Reaction:

$$C_2H_6O_2^+ \rightarrow C_2H_4^+ + H_2O$$

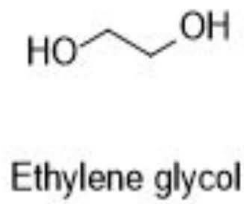
Summary of Expected Peaks:

- **m/z = 62 (M⁺)**: Molecular ion (may be weak or absent due to fragmentation).
- **m/z = 45**: Loss of one hydroxyl group (-OH).
- **m/z = 31**: Hydroxymethyl cation (CH_2OH^+) due to C–C bond cleavage.
- **m/z = 44**: Loss of water (H_2O).

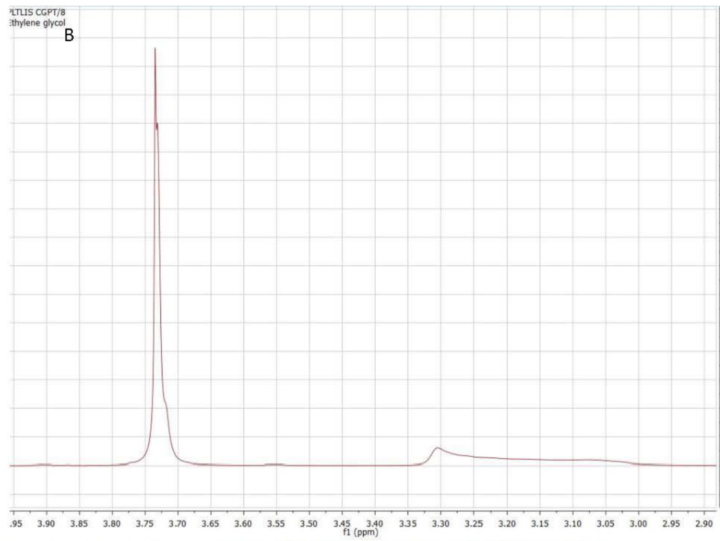
These fragmentation patterns provide a clear breakdown of **ethylene glycol** in mass spectrometry, helping identify key molecular features and confirm its structure.

Figure S11: (A) Chemical structure of Ethylene Glycol. (B) SDBS Database of Ethylene Glycol (11) (C) Textual output from ChatGPT-4 for prominent fragments of Ethylene Glycol spectra

A



B



C

For **ethylene glycol** ($HOCH_2CH_2OH$), the ^1H-NMR spectrum can be analyzed using the **n+1 rule.** Ethylene glycol has two hydroxyl groups attached to a two-carbon chain, and here's how its protons behave:

Structure:

Ethylene glycol consists of two hydroxyl groups (-OH) attached to a two-carbon chain ($HOCH_2CH_2OH$).

Proton Environments:

1. **Hydroxyl protons (-OH):**
   - The protons on the hydroxyl groups typically do not couple with the protons on the neighboring methylene groups (-CH_2-) because of rapid exchange with solvent.
   - These protons appear as **broad singlets.**
   - Chemical shift: Typically around 2-4 ppm (depends on solvent and hydrogen bonding).
2. **Methylene protons (-CH_2):**
   - There are two equivalent -CH_2 groups in ethylene glycol, both adjacent to a hydroxyl group.
   - Each -CH_2 group will experience splitting from the two equivalent protons on the neighboring -CH_2 group (2 protons), giving a 2+1 = 3 peaks (triplet).
   - Chemical shift: Typically around 3.5-4 ppm (due to proximity to electronegative oxygen atoms).

Summary of Proton Splitting and Chemical Shifts:

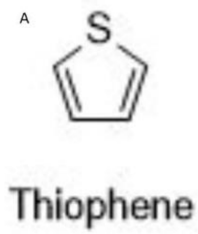
1. **Hydroxyl protons (-OH):**
   - Broad singlets, ~2-4 ppm.
2. **Methylene protons (-CH_2-):**
   - Triplets, ~3.5-4 ppm.

Overall Splitting Pattern:

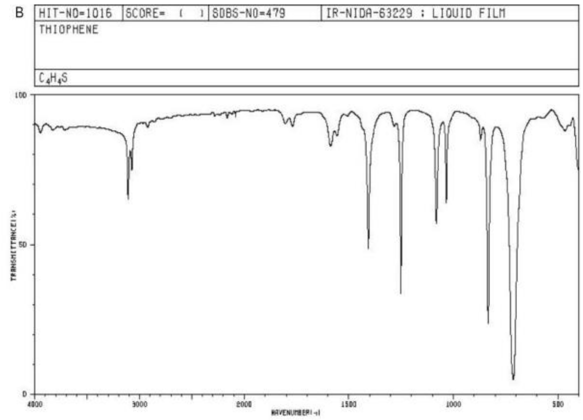
- -OH: Broad singlet.
- -CH_2: Triplet.

Figure S12: (A) Chemical structure of Ethylene Glycol. (B) Experimental spectra of Ethylene Glycol. (C) Textual output from ChatGPT-4 for prominent peaks of Ethylene Glycol spectra

A



B



C

For **thiophene (C_4H_4S)**, the IR spectrum will primarily reflect the **aromatic nature** of the thiophene ring, which contains a **sulfur atom** integrated into the five-membered ring.

Structure of Thiophene:

- **Aromatic heterocycle**: Thiophene has a five-membered ring with four carbon atoms and one sulfur atom.
- **C-H bonds**: Four C-H bonds attached to the carbon atoms in the ring.
- **C=C bonds**: Conjugated double bonds in the ring, as thiophene is an aromatic compound.
- **C-S bond**: The sulfur atom in the ring is bonded to two carbon atoms.

Key Peaks Expected in the IR Spectrum of Thiophene:

1. **C-H Stretch (Aromatic, 3000–3100 cm^{-1})**:
   - The C-H bonds on the aromatic ring produce **weak to moderate peaks** in the **3000–3100 cm^{-1}** range.
   - These peaks are characteristic of the stretching vibrations of C-H bonds in aromatic systems.
2. **C=C Stretch (Aromatic, 1450–1600 cm^{-1})**:
   - The **C=C bonds** in the conjugated ring system lead to **moderate to strong peaks** in the **1450–1600 cm^{-1}** region.
   - These peaks are typical of the stretching vibrations of C=C bonds in aromatic systems, with some variation due to the presence of the sulfur atom, which affects the electron density in the ring.
3. **C-S Stretch (600–750 cm^{-1})**:
   - The **C-S bond** in the thiophene ring leads to a peak in the **600–750 cm^{-1}** range.
   - This absorption is typically **strong** and appears as a sharp, well-defined peak.
   - Sulfur atoms in rings cause shifts in the C-S bond stretching, which is why the exact frequency is lower than typical C-H or C=C stretches.
4. **C-H Out-of-Plane Bending (700–900 cm^{-1})**:
   - The **out-of-plane bending** vibrations of the C-H bonds in the thiophene ring produce **moderate to strong peaks** in the **700–900 cm^{-1}** region.
   - For thiophene, these typically occur around **800–900 cm^{-1}**, indicating the monosubstitution pattern of the ring (one hydrogen per carbon atom).
5. **Ring Breathing Mode (1000–1100 cm^{-1})**:
   - The entire ring structure can undergo a **breathing vibration**, where all atoms move synchronously in and out.
   - This mode usually shows up in the **1000–1100 cm^{-1}** range as a **weak to moderate** peak.

Summary of Key Peaks:

- **C-H Stretch (Aromatic)**: 3000–3100 cm^{-1} (weak to moderate).
- **C=C Stretch (Aromatic)**: 1450–1600 cm^{-1} (moderate to strong).
- **C-S Stretch**: 600–750 cm^{-1} (strong).
- **C-H Out-of-Plane Bending**: 700–900 cm^{-1} (moderate to strong).
- **Ring Breathing Mode**: 1000–1100 cm^{-1} (weak to moderate).

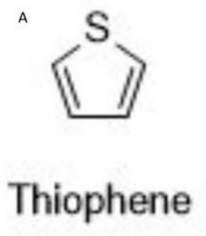
Additional Considerations:

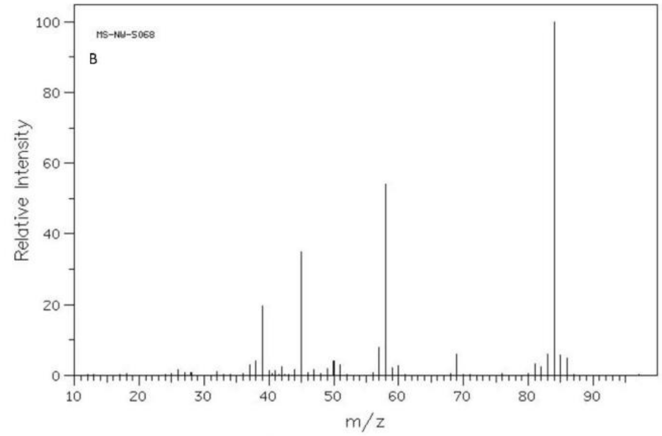
- **Aromaticity**: Thiophene is an aromatic compound, so the C-H and C=C stretches will resemble those of other aromatic systems like benzene, but the sulfur atom's presence shifts the positions slightly.
- **Substitution pattern**: The four C-H bonds on thiophene create a unique **out-of-plane bending pattern** around 800–900 cm^{-1}, which can be used to distinguish it from other aromatic compounds.
- **Electron-rich ring**: The sulfur atom contributes to the electron density in the ring, which affects the overall absorption intensities and shifts the C=C stretches slightly compared to pure hydrocarbons.

This analysis provides a detailed overview of the expected IR spectrum for thiophene. Let me know if you'd like further clarification!

Figure S13: (A) Chemical structure of Thiophene. (B) SDBS Database of Thiophene (11) (C) Textual output from ChatGPT-4 for prominent peaks of Thiophene spectra.

A





C

Mass Spectrometry Analysis for Thiophene (C_4H_4S)

Molecular Formula: C_4H_4S
Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 4 = 48.04
- Hydrogen (H): 1.008 × 4 = 4.032
- Sulfur (S): 32.06 × 1 = 32.06

Total = 84.13 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M⁺):

- The molecular ion peak, which corresponds to the intact molecule, appears at **m/z = 84**. This peak is typically strong for thiophene, indicating the complete molecule.

2. Cleavage of a CH Group (-CH):

- A common fragmentation pathway is the loss of one CH group (-CH, 13 Da), leading to a fragment at **m/z = 71**.

Reaction:

$$C_4H_4S^+ \rightarrow C_3H_3S^+ + CH$$

3. Loss of Two CH Groups (-2CH):

- Further fragmentation can result in the loss of two CH groups (-2CH, 26 Da), producing a fragment at **m/z = 58**.

Reaction:

$$C_4H_4S^+ \rightarrow C_2H_2S^+ + 2CH$$

4. Loss of the Sulfur Atom (S):

- The sulfur atom can also be cleaved from the molecule, resulting in a fragment at **m/z = 56**. This corresponds to the loss of 32 Da (the sulfur atom).

Reaction:

$$C_4H_4S^+ \rightarrow C_4H_4^+ + S$$

5. Cleavage of the Thiophene Ring:

- Thiophene's five-membered ring can undergo a variety of ring cleavages, producing a smaller fragment at **m/z = 39**, corresponding to a $C_3H_3^+$ fragment.

Reaction:

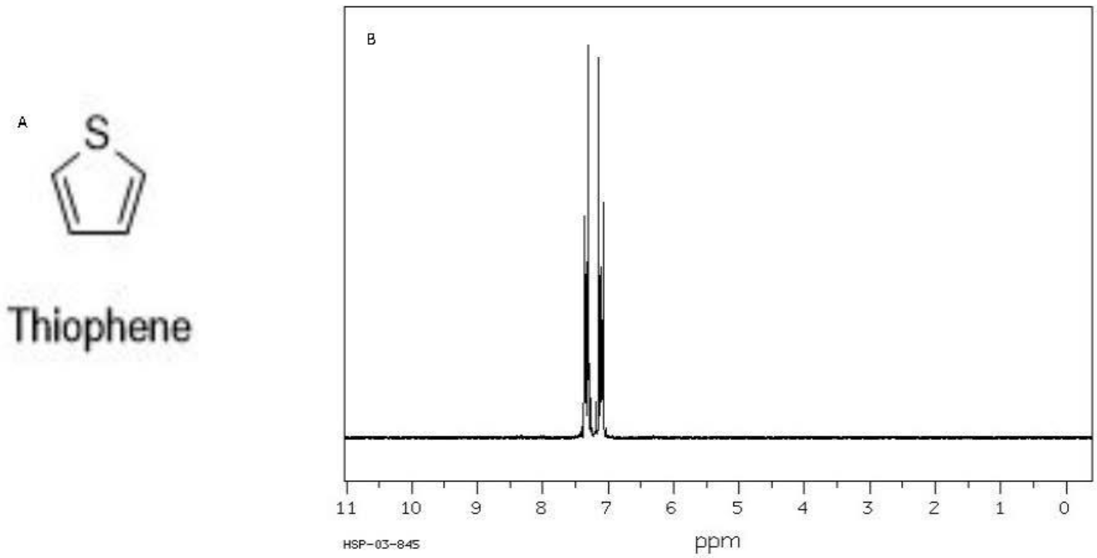
$$C_4H_4S^+ \rightarrow C_3H_3^+ + S$$

Summary of Expected Peaks:

- **m/z = 84 (M⁺)**: Molecular ion peak.
- **m/z = 71**: Loss of a CH group.
- **m/z = 58**: Loss of two CH groups.
- **m/z = 56**: Loss of the sulfur atom.
- **m/z = 39**: Fragment resulting from ring cleavage ($C_3H_3^+$).

These fragmentation pathways provide insight into the structure of **thiophene** and help identify it in mass spectrometry based on the characteristic peaks observed.

Figure S14: (A) Chemical structure of Thiophene. (B) SDBS Database of Thiophene (11) (C) Textual output from ChatGPT-4 for prominent fragments of Thiophene spectra.



C

For thiophene, which is a five-membered aromatic ring with four carbon atoms and one sulfur atom, we need to predict the proton splitting patterns using the **n+1 rule**. Thiophene is similar to benzene in terms of its aromatic system, but with sulfur replacing one of the carbons.

Structure Overview:

- **Thiophene ring:** The ring consists of four carbons and one sulfur.
- All five positions of the ring are aromatic.

Proton Splitting Patterns Using the n+1 Rule:

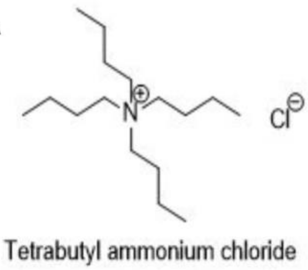
1. **Aromatic Protons:**
    - In thiophene, the four aromatic protons are equivalent in terms of their coupling, but their splitting pattern will reflect their interactions with each other.
    - The aromatic protons in thiophene can be analyzed based on their interactions with the neighboring protons on the ring. Each proton on the ring is adjacent to two other protons.
2. **Protons on the Ring:**
    - **Each Proton on the Ring:**
        - Each proton is adjacent to two other protons on the ring (one on either side).
        - **Using n+1 Rule:** Each proton sees 2 neighboring protons.
            - n = 2 (number of neighboring protons)
            - The result is a **triplet** for each proton.
    - **Chemical shift:** Aromatic protons in thiophene typically appear around **6.5 - 7.5 ppm**.

Summary Using the n+1 Rule:

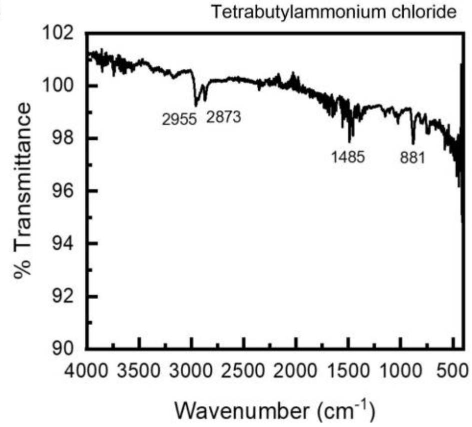
- **Aromatic Protons:**
    - Each proton on the thiophene ring will appear as a **triplet** around **6.5 - 7.5 ppm** (split by 2 neighboring protons on the ring).

Figure S15: (A) Chemical structure of Thiophene. (B) SDBS Database of Thiophene (11). (C) Textual output from ChatGPT-4 for prominent peaks of Thiophene spectra.

A



B



C

For **tetrabutylammonium chloride ($C_{16}H_{36}NCl$)**, the IR spectrum will mainly show signals related to the **alkyl chains (butyl groups)** and the **quaternary ammonium ion (N^+)**. The chloride ion (Cl^-) is not IR-active, so it will not show any distinct peaks.

Structure of Tetrabutylammonium Chloride:

- **Quaternary ammonium ion (N^+)**: Nitrogen bonded to four butyl groups, giving a positive charge.
- **Butyl groups ($-CH_2CH_2CH_2CH_3$)**: Four linear alkyl chains attached to the nitrogen atom.
- **Chloride ion (Cl^-)**: Paired with the ammonium ion as a counterion.

Key Peaks Expected in the IR Spectrum of Tetrabutylammonium Chloride:

1. **C-H Stretch (Alkyl, 2850–2960 cm^{-1})**:
   - The **C-H bonds** in the butyl chains will produce strong peaks in the **2850–2960 cm^{-1}** range.
   - These peaks correspond to the symmetric and asymmetric stretches of the **methylene (-CH_2-)** and **methyl (-CH_3)** groups.
   - You'll typically see:
     - Asymmetric stretching around **2920–2960 cm^{-1}**.
     - Symmetric stretching around **2850–2900 cm^{-1}**.
2. **C-H Bending (Methyl and Methylene, 1350–1480 cm^{-1})**:
   - The **bending vibrations** of the C-H bonds from both methyl and methylene groups will show **moderate to strong peaks** in the **1350–1480 cm^{-1}** range.
   - Expect:
     - **Methyl (-CH_3) deformation** (umbrella bending) near **1375 cm^{-1}**.
     - **Methylene (-CH_2) scissoring** around **1465 cm^{-1}**.
3. **N-CH_2 Stretch (950–1250 cm^{-1})**:
   - The stretching vibrations involving the **N-CH_2 bond** in the quaternary ammonium ion are expected to show peaks in the **950–1250 cm^{-1}** range.
   - This will typically appear as a **moderate to weak** peak, reflecting the presence of the nitrogen-carbon bond in the alkyl chains.
4. **C-N Stretch (Quaternary Ammonium, 1000–1100 cm^{-1})**:
   - The **C-N bond** in the quaternary ammonium ion shows a peak in the **1000–1100 cm^{-1}** range.
   - This peak is often **weak to moderate**, and it reflects the stretching of the carbon-nitrogen bonds in the tetrabutyl structure.
5. **CH_2 Rocking (700–800 cm^{-1})**:
   - The **rocking motion** of the methylene (-CH_2-) groups gives rise to a peak in the **700–800 cm^{-1}** range.
   - This is a weaker, lower-energy vibration but can still be **observable** in longer alkyl chains like the butyl groups.

Summary of Key Peaks:

- **C-H Stretch (Alkyl)**: 2850–2960 cm^{-1} (strong).
- **C-H Bending (Methyl & Methylene)**: 1350–1480 cm^{-1} (moderate to strong).
- **N-CH_2 Stretch**: 950–1250 cm^{-1} (moderate).
- **C-N Stretch**: 1000–1100 cm^{-1} (weak to moderate).
- **CH_2 Rocking**: 700–800 cm^{-1} (weak).

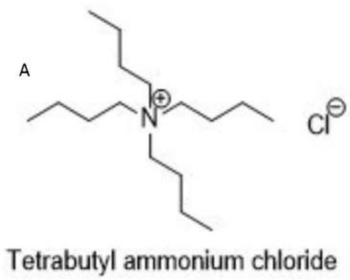
Additional Considerations:

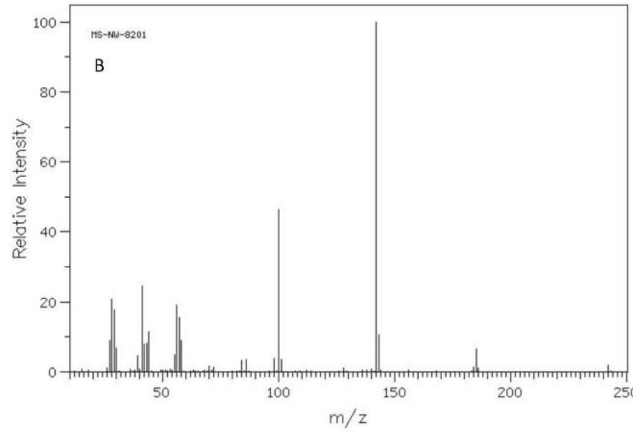
- **Alkyl dominance**: The spectrum of tetrabutylammonium chloride will be dominated by the alkyl chains' C-H stretching and bending vibrations, as the molecule is made up mostly of butyl groups.
- **Quaternary ammonium signature**: The C-N stretching vibrations will be weaker compared to the alkyl group signals, but they can still be distinguished around 1000–1100 cm^{-1}.
- **Chloride ion**: As mentioned, the chloride ion (Cl^-) does not contribute to IR-active vibrations, so no peaks will be seen related to the chloride itself.

This analysis covers the expected IR peaks for tetrabutylammonium chloride. Let me know if you'd like further details!

Figure S16: (A) Chemical structure of Tetrabutylammonium Chloride. (B) Experimental spectra of Tetrabutylammonium Chloride. (C) Textual output from ChatGPT-4 for prominent peaks of Tetrabutyl Ammonium Chloride spectra.

A





C

Mass Spectrometry Analysis for Tetrabutyl Ammonium Chloride ($C_{16}H_{36}NCl$)

Molecular Formula: $C_{16}H_{36}NCl$

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 16 = 192.16
- Hydrogen (H): 1.008 × 36 = 36.288
- Nitrogen (N): 14.01 × 1 = 14.01
- Chlorine (Cl): 35.45 × 1 = 35.45

Total = 277.91 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M^+):

- In mass spectrometry, **tetrabutyl ammonium chloride** typically shows fragmentation of the tetrabutyl ammonium ion itself rather than the chloride (Cl^-) anion. The molecular ion peak corresponding to $C_{16}H_{36}N^+$ (tetrabutyl ammonium) would appear at **m/z = 242**.
  However, the molecular ion is often weak or absent due to the instability of the large cation, and chloride does not typically appear directly in the positive ion spectrum.

2. Sequential Loss of Butyl Groups ($-C_4H_9$):

- A key fragmentation pathway involves the sequential loss of butyl groups ($-C_4H_9$, 57 Da each), resulting in a series of characteristic peaks:
  - Loss of one butyl group: **m/z = 185**.
  - Loss of two butyl groups: **m/z = 128**.
  - Loss of three butyl groups: **m/z = 71**.

  Reactions:

$$C_{16}H_{36}N^+ \rightarrow C_{12}H_{28}N^+ + C_4H_9 \quad (m/z = 185)$$
$$C_{12}H_{28}N^+ \rightarrow C_8H_{20}N^+ + C_4H_9 \quad (m/z = 128)$$
$$C_8H_{20}N^+ \rightarrow C_4H_{12}N^+ + C_4H_9 \quad (m/z = 71)$$

3. Cleavage of the Nitrogen Center:

- The nitrogen center in tetrabutyl ammonium can also undergo cleavage, resulting in fragments that correspond to the butyl group attached to nitrogen and a smaller cation. One of the smaller fragments might correspond to **m/z = 57**, which is a common fragment from the butyl group.

4. Presence of Chlorine (Isotopic Pattern):

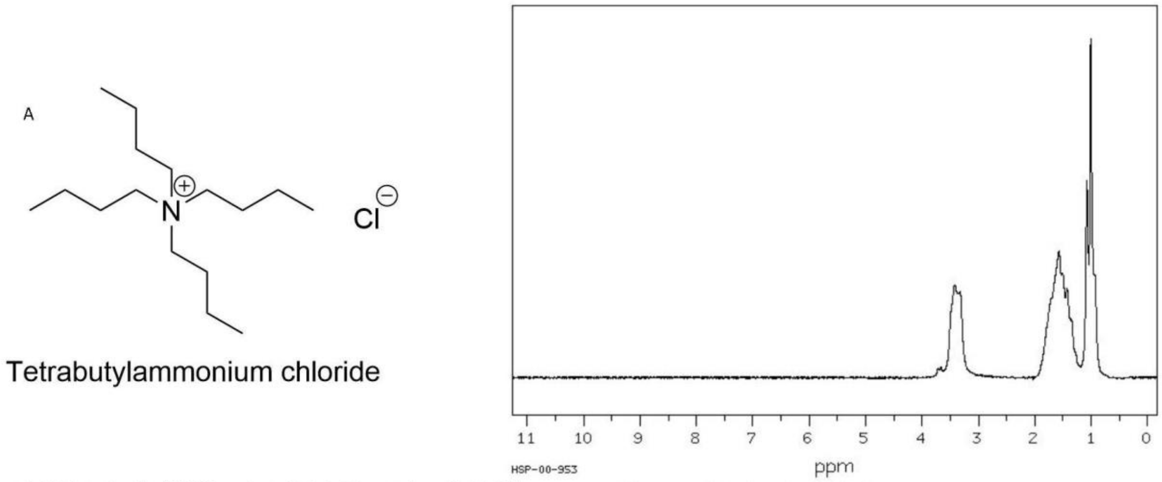
- Chlorine will manifest through its characteristic isotopic pattern (^{35}Cl and ^{37}Cl isotopes), though the chloride ion (Cl^-) is not typically visible in positive ion mass spectrometry. However, it will produce a minor isotopic pattern around **m/z = 35** and **m/z = 37** in negative ion mode, representing $^{35}Cl^-$ and $^{37}Cl^-$ respectively.

Summary of Expected Peaks:

- **m/z = 242 (M^+)**: Molecular ion (tetrabutyl ammonium cation).
- **m/z = 185**: Loss of one butyl group ($-C_4H_9$).
- **m/z = 128**: Loss of two butyl groups.
- **m/z = 71**: Loss of three butyl groups.
- **m/z = 57**: Butyl fragment.
- **m/z = 35 and 37**: Chlorine isotopic peaks (^{35}Cl and ^{37}Cl) seen in negative ion mode.

These fragmentation patterns help identify **tetrabutyl ammonium chloride** in mass spectrometry, particularly focusing on the sequential loss of the butyl groups from the tetrabutyl ammonium ion.

Figure S17: (A) Chemical structure of Tetrabutylammonium Chloride. (B) SDBS Database of Tetrabutylammonium Chloride (11). (C) Textual output from ChatGPT-4 for prominent fragments of Tetrabutyl Ammonium Chloride spectra.



C

No worries! Let's analyze the **1H NMR spectrum of tetrabutyl ammonium chloride.** This compound consists of a quaternary ammonium ion with four butyl groups attached to a central nitrogen, paired with a chloride counterion. The structure results in several distinct proton environments due to the four butyl chains.

Structure of Tetrabutyl Ammonium Chloride:

- **Chemical Formula**: $C_{16}H_{36}NCl$
- **Functional Groups**: Quaternary ammonium ion with four butyl groups.

Proton Environments:

1. **Methyl Protons ($-CH_3$) at the End of Each Butyl Chain:**
   - **Chemical Shift**: Typically resonates around 0.9 to 1.2 ppm. These protons are slightly shielded by the alkyl chain, resulting in an upfield shift.
   - **Splitting Pattern**: These protons appear as triplets due to coupling with the methylene protons adjacent to them in the butyl chain.
2. **Methylene Protons ($-CH_2-$) in the Butyl Chain:**
   - **Methylene Protons Adjacent to the Methyl Group**:
     - **Chemical Shift**: Around 1.25 to 1.6 ppm.
     - **Splitting Pattern**: Typically a multiplet, often appearing as a quintet due to coupling with both the neighboring methylene protons and the terminal methyl protons.
   - **Central Methylene Protons**:
     - **Chemical Shift**: These protons generally resonate from 1.6 to 2.0 ppm as they are further deshielded by their distance from the terminal methyl groups.
     - **Splitting Pattern**: Appear as complex multiplets, influenced by the long alkyl chain.
3. **Methine Protons (–CH–) Attached to the Nitrogen Atom:**
   - **Chemical Shift**: These protons are deshielded by the electron-withdrawing effect of the positively charged nitrogen, typically appearing around 3.0 to 3.5 ppm.
   - **Splitting Pattern**: Typically a multiplet, due to coupling with adjacent methylene protons.

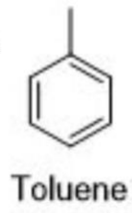
Summary of Splitting Patterns and Chemical Shifts:

- **Methyl Protons at Chain End**: Triplet, ~0.9-1.2 ppm.
- **Methylene Protons Adjacent to Methyl Groups**: Quintet, ~1.25-1.6 ppm.
- **Central Methylene Protons**: Multiplet, ~1.6-2.0 ppm.
- **Methine Protons Attached to Nitrogen**: Multiplet, ~3.0-3.5 ppm.

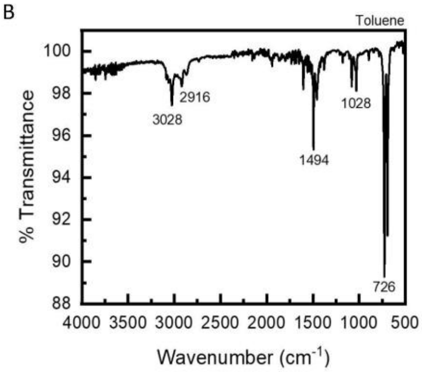
This detailed description helps in understanding the 1H NMR spectrum of tetrabutyl ammonium chloride, illustrating how the arrangement of butyl groups and their relation to the central nitrogen influence their chemical shifts and splitting patterns. This spectrum is characteristic of long alkyl chains and is useful for identifying such structures in a compound.

Figure S18: (A) Chemical structure of Tetrabutylammonium Chloride. (B) SDBS Database of Tetrabutylammonium Chloride (11). (C) Textual output from ChatGPT-4 for prominent peaks of Tetrabutylammonium Chloride spectra.

A



B



C

For **toluene (C_7H_8)**, the IR spectrum will primarily show peaks related to the **aromatic ring** and the **methyl group (-CH_3)** attached to the benzene ring.

Structure of Toluene:

- **Aromatic ring**: A benzene ring (C_6H_5) with alternating single and double bonds between carbon atoms.
- **Methyl group (-CH_3)**: Attached to one of the carbons on the benzene ring.

Key Peaks Expected in the IR Spectrum of Toluene:

1. **C-H Stretch (Aromatic, 3000–3100 cm^{-1})**:
   - The C-H bonds in the **aromatic ring** produce **weak to moderate peaks** in the **3000–3100 cm^{-1}** range.
   - These peaks are characteristic of **aromatic C-H stretching**.
2. **C-H Stretch (Methyl, 2850–2960 cm^{-1})**:
   - The **methyl group (-CH_3)** attached to the ring will show **moderate to strong peaks** due to **symmetric and asymmetric C-H stretching** in the **2850–2960 cm^{-1}** range.
     - **Asymmetric stretching** around **2920–2960 cm^{-1}**.
     - **Symmetric stretching** around **2850–2900 cm^{-1}**.
3. **C=C Stretch (Aromatic, 1450–1600 cm^{-1})**:
   - The **C=C stretching** vibrations of the benzene ring produce **weak to moderate peaks** in the **1450–1600 cm^{-1}** region.
   - These correspond to the **stretching of conjugated double bonds** in the ring and will appear as multiple peaks due to different resonance modes in the aromatic system.
4. **C-H Bending (Methyl, 1350–1470 cm^{-1})**:
   - The bending vibrations of the methyl (-CH_3) group will give rise to **moderate to strong peaks** in the **1350–1470 cm^{-1}** range.
     - **Methyl bending** (deformation) occurs near **1375 cm^{-1}**, a characteristic strong band for the -CH_3 group.
5. **C-H Out-of-Plane Bending (Aromatic, 700–900 cm^{-1})**:
   - The **out-of-plane bending** of the C-H bonds in the aromatic ring produces **strong characteristic peaks** in the **700–900 cm^{-1}** range.
   - For toluene, the **monosubstitution pattern** of the aromatic ring (one substituent, the methyl group) leads to a distinct peak around **730–770 cm^{-1}**, typical of monosubstituted benzenes.
6. **Ring Breathing Mode (Around 1000 cm^{-1})**:
   - The entire benzene ring undergoes a **breathing mode vibration**, which produces a **weak peak** near **1000 cm^{-1}**.

Summary of Key Peaks:

- **C-H Stretch (Aromatic)**: 3000–3100 cm^{-1} (weak to moderate).
- **C-H Stretch (Methyl)**: 2850–2960 cm^{-1} (moderate to strong).
- **C=C Stretch (Aromatic)**: 1450–1600 cm^{-1} (weak to moderate).
- **C-H Bending (Methyl)**: 1350–1470 cm^{-1} (moderate to strong, around 1375 cm^{-1}).
- **C-H Out-of-Plane Bending (Aromatic)**: 700–900 cm^{-1} (strong, particularly around 730–770 cm^{-1}).
- **Ring Breathing Mode**: Around 1000 cm^{-1} (weak).

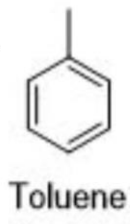
Additional Considerations:

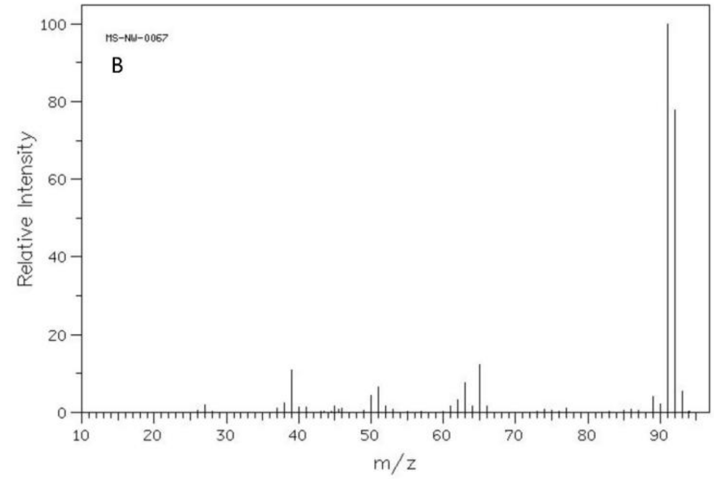
- **Aromatic ring**: The C=C stretching and out-of-plane bending peaks are typical for aromatic compounds and are used to identify the benzene ring.
- **Methyl group**: The -CH_3 group's strong bending vibration at around 1375 cm^{-1} is a good indicator of the presence of a methyl group attached to an aromatic system.
- **Monosubstitution**: The pattern of peaks in the **700–900 cm^{-1}** range can help confirm that the ring is monosubstituted (i.e., one group attached, in this case, the methyl group).

This breakdown provides a detailed overview of the IR spectrum for toluene. Let me know if you need further clarification!

Figure S19: (A) Chemical structure of Toluene. (B) Experimental spectra of Toluene. (C) Textual output from ChatGPT-4 for prominent peaks of Toluene spectra.

A





C

Mass Spectrometry Analysis for Toluene (C_7H_8)

Molecular Formula: C_7H_8

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 7 = 84.07
- Hydrogen (H): 1.008 × 8 = 8.064

Total = 92.13 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M^+):

- The molecular ion peak for **toluene** corresponds to the intact molecule and appears at **m/z = 92**. This is the mass of the complete molecule, C_7H_8, and it's usually a prominent peak in the spectrum.

2. Loss of a Methyl Group ($-CH_3$):

- A common fragmentation pathway is the loss of the methyl group ($-CH_3$, 15 Da), resulting in a benzyl cation ($C_6H_5^+$) at **m/z = 77**. This is one of the most characteristic fragments for toluene, as the benzyl cation is particularly stable.

  Reaction:

$$C_7H_8^+ \rightarrow C_6H_5^+ + CH_3 \quad (m/z = 77)$$

3. Formation of the Tropylium Ion:

- The toluene molecule can also rearrange to form a tropylium ion ($C_7H_7^+$), which is a highly stable structure. This produces a peak at **m/z = 91**. The tropylium ion is often one of the strongest peaks in the spectrum of toluene.

  Reaction:

$$C_7H_8^+ \rightarrow C_7H_7^+ + H$$

4. Further Fragmentation of the Benzyl Cation:

- The benzyl cation (m/z = 77) can undergo further fragmentation, leading to the loss of a hydrogen atom (H), resulting in a fragment at **m/z = 76**. This peak is less intense compared to the benzyl cation itself but can still appear in the spectrum.

5. Cleavage of the Benzene Ring:

- Toluene can undergo cleavage of the benzene ring, although these fragments are less common. Cleavage can result in smaller fragments like **m/z = 65** ($C_5H_5^+$) and **m/z = 51** ($C_4H_3^+$), which represent partial breakdown of the aromatic ring structure.

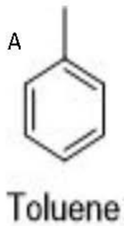
Summary of Expected Peaks:

- **m/z = 92 (M^+)**: Molecular ion.
- **m/z = 91**: Formation of the stable tropylium ion ($C_7H_7^+$).
- **m/z = 77**: Loss of the methyl group, forming the benzyl cation ($C_6H_5^+$).
- **m/z = 76**: Further fragmentation of the benzyl cation.
- **m/z = 65**: Partial cleavage of the benzene ring.
- **m/z = 51**: Further breakdown of the aromatic ring.

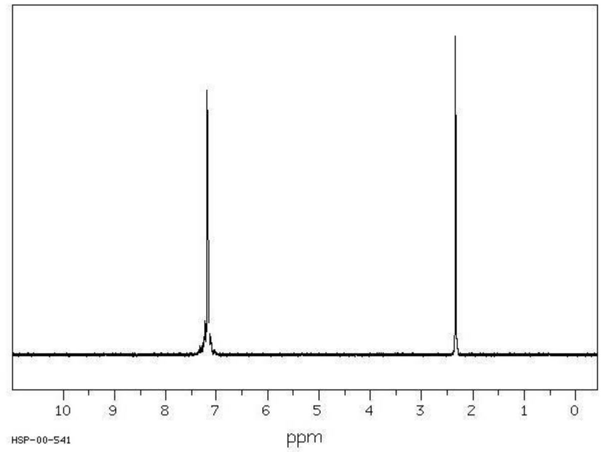
These fragmentation pathways provide a clear picture of the breakdown of **toluene** in mass spectrometry, with the **m/z = 91** and **m/z = 77** peaks being particularly diagnostic for this compound.

Figure S20: (A) Chemical structure of Toluene. (B) SDBS Database of Toluene (11). (C) Textual output from ChatGPT-4 for prominent fragments of Toluene spectra.

A



B



C

For **toluene** (methylbenzene), the molecule consists of a benzene ring with a methyl group ($-CH_3$) attached. Let's analyze the ^1H NMR spectrum of toluene using the n+1 rule to determine the splitting patterns.

Structure Overview:

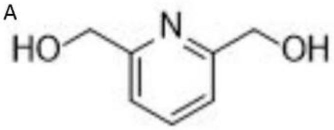
- Benzene ring with a **methyl group** at the para position.
- Aromatic protons on the ring and the protons of the methyl group.

Proton Splitting Patterns Using the n+1 Rule:

1. **Methyl Group Protons ($-CH_3$):**
   - The protons on the methyl group are adjacent to the aromatic ring.
   - **Using n+1 Rule:** Each methyl proton is coupled with the adjacent aromatic protons.
     - The methyl group has three protons.
     - These protons are split by the adjacent methylene protons on the aromatic ring.
     - The result is a **singlet** because the methyl group is only adjacent to the ring and there is no additional splitting beyond the ring.
   - **Chemical shift:** Methyl protons typically appear around **2.1 - 2.5 ppm**.
2. **Aromatic Protons:**
   - The benzene ring has six protons that are symmetrically distributed. In toluene, one of the positions is occupied by the methyl group, affecting the splitting pattern of the remaining five protons.
   - **Protons at the Positions Ortho to Methyl Group** (2 and 6 positions relative to the methyl group):
     - These protons are coupled with the adjacent ortho proton and the meta proton.
     - **Using n+1 Rule:** Each ortho proton is split by 2 neighboring protons (the other ortho proton and one meta proton).
       - n = 2 (number of neighboring protons)
       - The result is a **doublet of doublets (dd)**.
     - **Chemical shift:** Typically around **7.0 - 7.5 ppm**.
   - **Protons at the Meta Positions** (3 and 5 positions relative to the methyl group):
     - These protons are coupled with two ortho protons.
     - **Using n+1 Rule:** Each meta proton is split by 2 neighboring ortho protons.
       - n = 2 (number of neighboring protons)
       - The result is a **triplet**.
     - **Chemical shift:** Typically around **7.0 - 7.5 ppm**.
   - **Proton at the Para Position** (relative to the methyl group):
     - This proton is coupled with the two adjacent protons (one at each side).
     - **Using n+1 Rule:** The para proton is typically part of the overall aromatic multiplet, and its exact pattern is influenced by its symmetrical coupling with the rest of the ring.
     - **Chemical shift:** Typically around **7.0 - 7.5 ppm**.

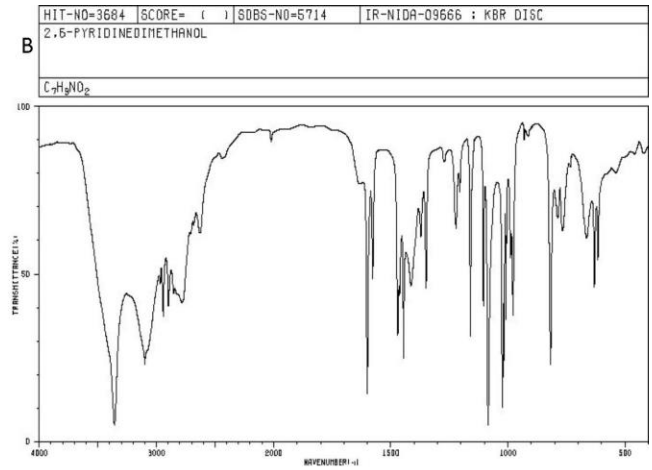
Figure S21: (A) Chemical structure of Toluene. (B) SDBS Database of Toluene (11). (C) Textual output from ChatGPT-4 for prominent peaks of Toluene spectra.

A



2,6-pyridinedimethanol

B



C

For **2,6-pyridinedimethanol ($C_7H_9NO_2$)**, the IR spectrum will reflect the features of the **pyridine ring**, the **hydroxyl (-OH) groups** attached to the methanol moieties at positions 2 and 6, and the **C-H bonds** from the methylene ($-CH_2-$) groups.

Structure of 2,6-Pyridinedimethanol:

- **Pyridine ring**: A six-membered ring with alternating single and double bonds, and a nitrogen atom at position 1.
- **Hydroxymethyl groups ($-CH_2OH$)**: Attached at the 2 and 6 positions of the pyridine ring.

Key Peaks Expected in the IR Spectrum of 2,6-Pyridinedimethanol:

1. **O-H Stretch (Hydroxyl, 3200–3600 cm^{-1})**:
   - The **hydroxyl (-OH)** groups from the alcohol functional groups at positions 2 and 6 will show **broad, strong absorption** between **3200 and 3600 cm^{-1}**.
   - This broad peak is characteristic of alcohols and is caused by hydrogen bonding, which can shift and broaden the peak further due to intermolecular interactions.
2. **C-H Stretch (Alkyl, 2850–2950 cm^{-1})**:
   - The C-H stretching from the **methylene ($-CH_2-$)** groups in the hydroxymethyl moieties will show **moderate peaks** in the **2850–2950 cm^{-1}** range.
   - These will be typical symmetric and asymmetric stretching vibrations of alkyl C-H bonds.
3. **C-H Stretch (Aromatic, 3000–3100 cm^{-1})**:
   - The aromatic C-H bonds in the pyridine ring will give **weak to moderate peaks** in the **3000–3100 cm^{-1}** region.
   - This is characteristic of aromatic systems, with lower intensity than aliphatic C-H stretches.
4. **C-O Stretch (Alcohol, 1000–1100 cm^{-1})**:
   - The **C-O stretching vibration** from the hydroxyl groups will produce a peak in the **1000–1100 cm^{-1}** range.
   - This peak is usually **strong** and well-defined, indicative of the presence of alcohol functional groups.
5. **C=C Stretch (Aromatic, 1450–1600 cm^{-1})**:
   - The **C=C bonds** in the pyridine ring will produce **weak to moderate peaks** in the **1450–1600 cm^{-1}** range.
   - These are typical of the stretching vibrations in aromatic rings.
6. **C-H Bending (Alkyl, 1350–1480 cm^{-1})**:
   - The **bending (scissoring) vibrations** of the methylene ($-CH_2-$) groups will result in **moderate peaks** in the **1350–1480 cm^{-1}** range.
   - This peak is due to the bending motions of the C-H bonds in the alkyl chain.
7. **C-N Stretch (Pyridine, 1100–1250 cm^{-1})**:
   - The **C-N stretching** vibration from the nitrogen in the pyridine ring will appear in the **1100–1250 cm^{-1}** range.
   - This peak may be weak to moderate in intensity and is typical of nitrogen-containing heterocycles like pyridine.
8. **O-H Bending (Alcohol, 1350–1450 cm^{-1})**:
   - There may be a weaker **O-H bending** vibration due to the hydroxyl groups in the **1350–1450 cm^{-1}** range.
9. **C-H Out-of-Plane Bending (Aromatic, 750–900 cm^{-1})**:
   - The **out-of-plane bending** vibrations of the C-H bonds in the aromatic ring will produce peaks in the **750–900 cm^{-1}** range, typical of aromatic compounds.

Summary of Key Peaks:

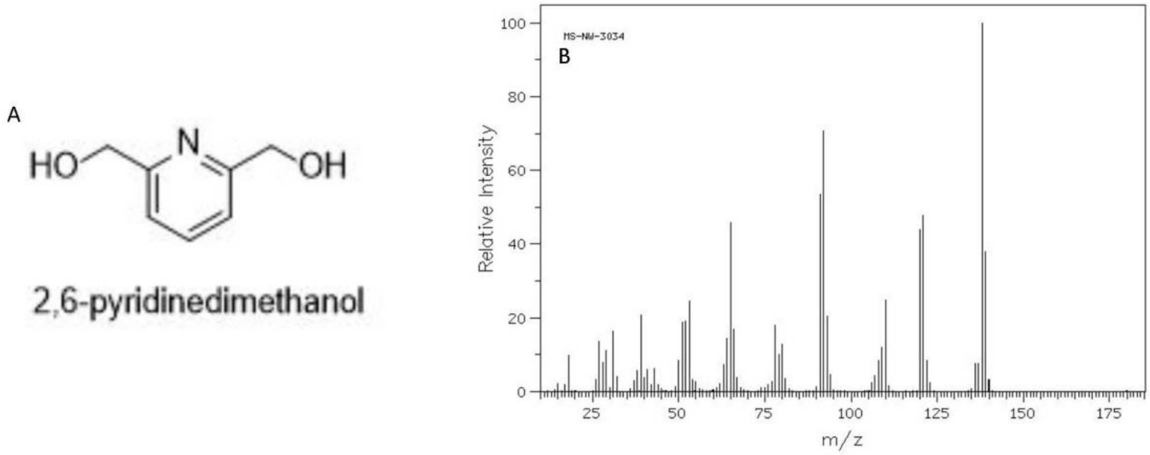
- **O-H Stretch (Alcohol)**: 3200–3600 cm^{-1} (broad, strong).
- **C-H Stretch (Alkyl)**: 2850–2950 cm^{-1} (moderate).
- **C-H Stretch (Aromatic)**: 3000–3100 cm^{-1} (weak to moderate).
- **C-O Stretch (Alcohol)**: 1000–1100 cm^{-1} (strong).
- **C=C Stretch (Aromatic)**: 1450–1600 cm^{-1} (moderate).
- **C-H Bending (Alkyl)**: 1350–1480 cm^{-1} (moderate).
- **C-N Stretch (Pyridine)**: 1100–1250 cm^{-1} (weak to moderate).
- **O-H Bending (Alcohol)**: 1350–1450 cm^{-1} (weak).
- **C-H Out-of-Plane Bending (Aromatic)**: 750–900 cm^{-1} (weak to moderate).

Additional Considerations:

- **Hydroxyl groups**: The two hydroxymethyl groups at positions 2 and 6 will cause strong, broad absorption in the O-H stretch region due to hydrogen bonding.
- **Aromatic ring**: The aromatic ring's C-H stretches, and C=C stretches will resemble those seen in pyridine derivatives, with the nitrogen atom influencing the electronic distribution and peak positions.
- **Aliphatic and aromatic distinctions**: The presence of both alkyl and aromatic groups will result in clear separation between their respective C-H stretch and bending regions.

This profile gives a detailed breakdown of the expected IR spectrum for 2,6-pyridinedimethanol. Let me know if you need more details!

Figure S22: (A) Chemical structure of 2,6-Pyridinedimethanol (B) SDBS Database of 2,6-Pyridinedimethanol (11). (C) Textual output from ChatGPT-4 for prominent peaks of 2,6-Pyridinedimethanol spectra.



C

Mass Spectrometry Analysis for 2,6-Pyridinedimethanol ($C_7H_9NO_2$)

Molecular Formula: $C_7H_9NO_2$

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 7 = 84.07
- Hydrogen (H): 1.008 × 9 = 9.072
- Nitrogen (N): 14.01 × 1 = 14.01
- Oxygen (O): 16.00 × 2 = 32.00

Total = 139.15 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M⁺):

- The molecular ion peak for **2,6-Pyridinedimethanol** corresponds to the intact molecule, appearing at **m/z = 139**. This peak represents the full molecule with no fragmentation.

2. Loss of Hydroxyl Group (-OH):

- One of the most common fragmentation pathways is the loss of a hydroxyl group (-OH, 17 Da). This results in a fragment at **m/z = 122**, which represents the remaining part of the molecule after the loss of one hydroxyl group.

  Reaction:

$$C_7H_9NO_2^+ \rightarrow C_7H_8NO^+ + OH \quad (m/z = 122)$$

3. Loss of Formaldehyde (CH_2O):

- Another key fragmentation involves the loss of formaldehyde (CH_2O, 30 Da), which originates from one of the hydroxymethyl groups attached to the pyridine ring. This leads to a fragment at **m/z = 109**.

  Reaction:

$$C_7H_9NO_2^+ \rightarrow C_6H_7NO^+ + CH_2O \quad (m/z = 109)$$

4. Loss of Both Hydroxyl Groups (-2OH):

- Sequential loss of both hydroxyl groups (-OH × 2 = 34 Da) results in a fragment at **m/z = 105**, which represents the remaining pyridine ring with attached alkyl chains.

  Reaction:

$$C_7H_9NO_2^+ \rightarrow C_7H_7N^+ + 2OH \quad (m/z = 105)$$

5. Cleavage of Pyridine Ring:

- The pyridine ring can also undergo fragmentation. This can produce a fragment at **m/z = 79**, corresponding to the pyridyl cation ($C_5H_5N^+$), which is a common feature in the mass spectra of pyridine derivatives.

  Reaction:

$$C_7H_9NO_2^+ \rightarrow C_5H_5N^+ + C_2H_4O_2 \quad (m/z = 79)$$

6. Cleavage of Methanol Group (-CH_2OH):

- Loss of a single methanol group (-CH_2OH, 31 Da) can occur, leading to a fragment at **m/z = 108**. This corresponds to the molecule minus one methanol group.

  Reaction:

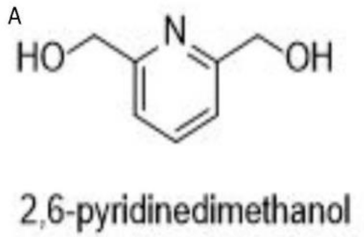
$$C_7H_9NO_2^+ \rightarrow C_6H_7NO^+ + CH_2OH \quad (m/z = 108)$$

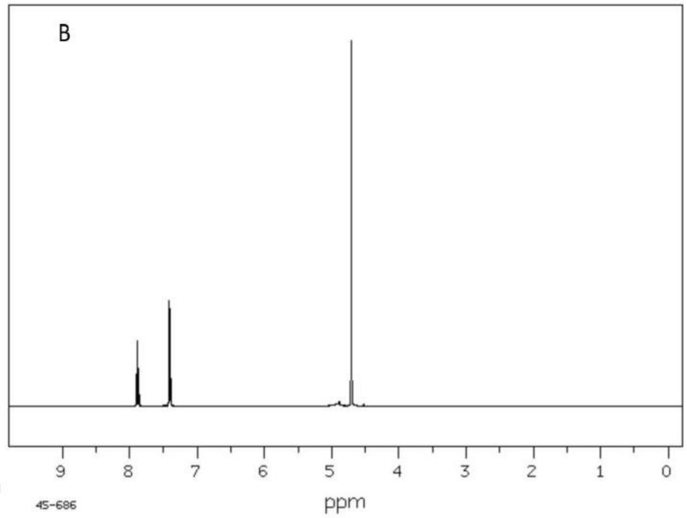
Summary of Expected Peaks:

- **m/z = 139 (M⁺)**: Molecular ion peak.
- **m/z = 122**: Loss of one hydroxyl group (-OH).
- **m/z = 109**: Loss of formaldehyde (CH_2O).
- **m/z = 105**: Loss of both hydroxyl groups (-2OH).
- **m/z = 108**: Loss of one methanol group (-CH_2OH).
- **m/z = 79**: Pyridyl cation ($C_5H_5N^+$) from cleavage of the pyridine ring.

These fragmentation pathways provide a detailed breakdown of **2,6-Pyridinedimethanol** in mass spectrometry, with the key peaks representing the stepwise loss of functional groups and cleavage of the pyridine ring.

Figure S23: (A) Chemical structure of 2,6-Pyridinedimethanol (B) SDBS Database of 2,6-Pyridinedimethanol (11). (C) Textual output from ChatGPT-4 for prominent fragments of 2,6-Pyridinedimethanol spectra.





C

For 2,6-Pyridinedimethanol, the structure consists of a pyridine ring with two hydroxymethyl ($-CH_2OH$) groups at the 2- and 6-positions. Let's apply the n+1 rule to predict the proton splitting patterns in its ^1H NMR spectrum.

Structure Overview:

- Pyridine ring with hydroxymethyl groups at the 2- and 6-positions.
- Aromatic protons on the pyridine ring (at positions 3, 4, and 5).
- Hydroxymethyl ($-CH_2OH$) groups at positions 2 and 6.

Proton Splitting Patterns Using the n+1 Rule:

1. **Hydroxymethyl Protons ($-CH_2OH$):**
    - The methylene protons ($-CH_2-$) are attached to the pyridine ring at the 2- and 6-positions.
    - **Using n+1 Rule:** Each $-CH_2OH$ group is adjacent to the ring and not directly coupled with the other $-CH_2OH$ group (due to the distance).
        - Each methylene proton is split by its neighboring hydroxyl proton (-OH), but the -OH proton typically shows up as a broad singlet due to exchange with water in most solvents.
        - **Methylene protons:** Appear as a singlet due to no coupling with other protons on the ring.
    - **Chemical shift:** Methylene protons ($-CH_2-$) typically appear around **4.5 - 5.0 ppm.**
    - **Hydroxyl protons (-OH):** If visible, these appear as a **broad singlet** around **2.0 - 4.0 ppm.**
2. **Aromatic Protons (positions 3, 4, and 5):**
    - These protons are on the pyridine ring and are split by adjacent protons.
    - **Position 3 Proton:**
        - This proton is adjacent to the proton at position 4.
        - **Using n+1 Rule:** The proton at position 3 is split by 1 proton (the proton at position 4).
        - n = 1 (number of neighboring protons)
        - The result is a **doublet.**
        - **Chemical shift:** Appears around **7.0 - 8.0 ppm.**
    - **Position 4 Proton:**
        - This proton is adjacent to the protons at positions 3 and 5.
        - **Using n+1 Rule:** The proton at position 4 is split by 2 neighboring protons (positions 3 and 5).
        - n = 2 (number of neighboring protons)
        - The result is a **triplet.**
        - **Chemical shift:** Appears around **7.0 - 8.0 ppm.**
    - **Position 5 Proton:**
        - This proton is adjacent to the proton at position 4.
        - **Using n+1 Rule:** The proton at position 5 is split by 1 proton (the proton at position 4).
        - n = 1 (number of neighboring protons)
        - The result is a **doublet.**
        - **Chemical shift:** Appears around **7.0 - 8.0 ppm.**

Summary Using the n+1 Rule:

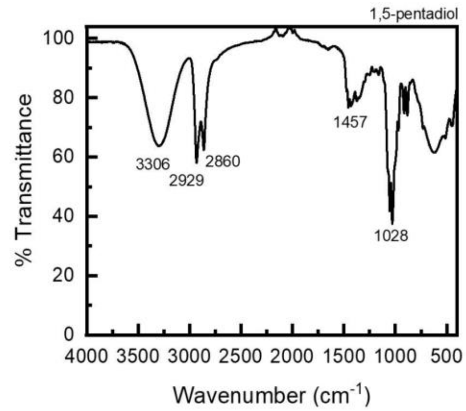
- **Hydroxymethyl Protons ($-CH_2OH$):**
    - **Singlet at 4.5 - 5.0 ppm** (no coupling with neighboring protons).
- **Aromatic Protons:**
    - **Doublet at 7.0 - 8.0 ppm** (position 3, split by 1 proton at position 4).
    - **Triplet at 7.0 - 8.0 ppm** (position 4, split by 2 neighboring protons at positions 3 and 5).
    - **Doublet at 7.0 - 8.0 ppm** (position 5, split by 1 proton at position 4).

Figure S24: (A) Chemical structure of 2,6-Pyridinedimethanol (B) SDBS Database of 2,6-Pyridinedimethanol (11). (C) Textual output from ChatGPT-4 for prominent peaks of 2,6-Pyridinedimethanol spectra.

A

1,5-pentanediol

B



C

For **1,5-pentanediol ($C_5H_{12}O_2$)**, the IR spectrum will reflect the presence of two **hydroxyl (-OH)** groups and the **alkyl chain** of five carbon atoms.

Structure of 1,5-Pentanediol:

- **Aliphatic chain**: A straight-chain hydrocarbon consisting of five carbon atoms.
- **Hydroxyl groups (-OH)**: Two hydroxyl groups attached to the 1st and 5th carbon atoms.

Key Peaks Expected in the IR Spectrum of 1,5-Pentanediol:

1. **O-H Stretch (Broad, 3200–3600 cm⁻¹):**
   - The two **hydroxyl (-OH)** groups produce a **broad, strong absorption band** in the **3200–3600 cm⁻¹** region.
   - This broadening is typical of alcohols due to **hydrogen bonding**, which can lead to increased peak width.
2. **C-H Stretch (Alkyl, 2850–2960 cm⁻¹):**
   - The C-H stretching from the methylene ($-CH_2-$) and methyl ($-CH_3$) groups will produce **strong peaks** in the **2850–2960 cm⁻¹** range.
   - This includes:
     - **Asymmetric stretching** around **2920–2960 cm⁻¹**.
     - **Symmetric stretching** around **2850–2900 cm⁻¹**.
3. **O-H Bending (Alcohol, 1400–1470 cm⁻¹):**
   - There may be a weaker, yet noticeable peak due to **O-H bending (scissoring)** in the **1400–1470 cm⁻¹** range.
   - This corresponds to the bending of the hydroxyl groups.
4. **C-O Stretch (Alcohol, 1050–1150 cm⁻¹):**
   - The **C-O stretching vibrations** from the hydroxyl groups will produce **strong peaks** in the **1050–1150 cm⁻¹** range.
   - This peak is typically sharp and intense due to the presence of two hydroxyl groups.
5. **C-H Bending (Methylene and Methyl, 1350–1470 cm⁻¹):**
   - The **bending vibrations** of the methylene ($-CH_2-$) and methyl ($-CH_3$) groups produce **moderate peaks** in the **1350–1470 cm⁻¹** range.
   - These are related to C-H scissoring and deformation.
6. **C-H Rocking (Methylene, 700–900 cm⁻¹):**
   - The **rocking motion** of the methylene ($-CH_2-$) groups gives rise to **moderate peaks** in the **700–900 cm⁻¹** range.
   - This is indicative of the longer aliphatic chain.

Summary of Key Peaks:

- **O-H Stretch (Hydroxyl)**: Broad, strong, 3200–3600 cm⁻¹.
- **C-H Stretch (Alkyl)**: 2850–2960 cm⁻¹ (strong).
- **O-H Bending**: 1400–1470 cm⁻¹ (weak to moderate).
- **C-O Stretch**: 1050–1150 cm⁻¹ (strong).
- **C-H Bending (Methylene & Methyl)**: 1350–1470 cm⁻¹ (moderate).
- **C-H Rocking (Methylene)**: 700–900 cm⁻¹ (moderate).

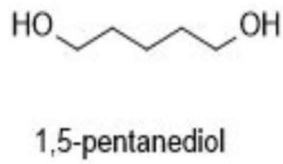
Additional Considerations:

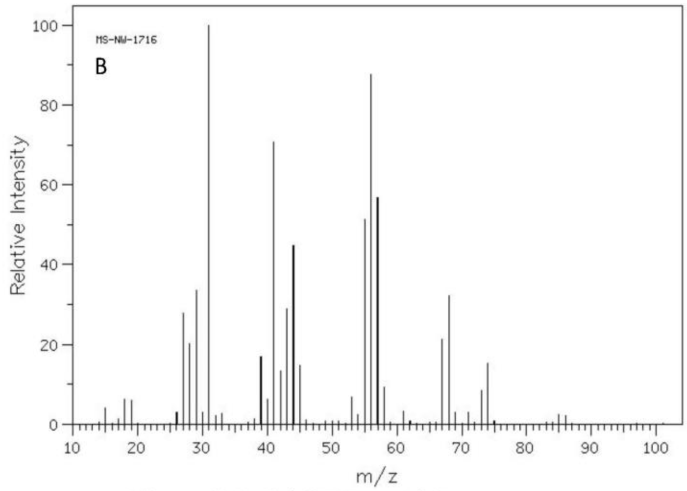
- **Hydrogen bonding**: The extensive hydrogen bonding in 1,5-pentanediol can lead to broader O-H stretching peaks compared to other simple alcohols.
- **Symmetrical structure**: The symmetrical nature of the molecule contributes to clearer absorption features, particularly in the C-H stretching region.

This profile provides a comprehensive overview of the expected IR spectrum for 1,5-pentanediol. Let me know if you need any further information!

Figure S25: (A) Chemical structure of 1,5-pentanediol. (B) Experimental spectra of 1,5-pentanediol. (C) Textual output from ChatGPT-4 for prominent peaks of 1,5-pentanediol spectra.

A





Mass Spectrometry Analysis for 1,5-Pentanediol ($C_5H_{12}O_2$)

Molecular Formula: $C_5H_{12}O_2$

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 5 = 60.05
- Hydrogen (H): 1.008 × 12 = 12.096
- Oxygen (O): 16.00 × 2 = 32.00

Total = 104.15 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M^+):

- The molecular ion peak for **1,5-Pentanediol** corresponds to the intact molecule, appearing at **m/z = 104**. This peak represents the entire molecule with no fragmentation.

2. Loss of Hydroxyl Group (-OH):

- A common fragmentation pathway involves the loss of a hydroxyl group (-OH, 17 Da), resulting in a fragment at **m/z = 87**. This corresponds to the molecule with one hydroxyl group removed.

  Reaction:

$$\mathrm{C_5H_{12}O_2^+ \rightarrow C_5H_{11}O^+ + OH} \quad (m/z = 87)$$

3. Loss of Water (H_2O):

- Another key fragmentation involves the loss of a water molecule (H_2O, 18 Da), which is typical for diols. This results in a fragment at **m/z = 86**.

  Reaction:

$$\mathrm{C_5H_{12}O_2^+ \rightarrow C_5H_{10}^+ + H_2O} \quad (m/z = 86)$$

4. Loss of Both Hydroxyl Groups (-2OH):

- The sequential loss of both hydroxyl groups (-OH × 2 = 34 Da) leads to a fragment at **m/z = 70**. This is the base fragment for the remaining hydrocarbon chain of 1,5-pentanediol.

  Reaction:

$$\mathrm{C_5H_{12}O_2^+ \rightarrow C_5H_{10}^+ + 2OH} \quad (m/z = 70)$$

5. Cleavage of Carbon Chain (C_5H_{10} Fragmentation):

- The cleavage of the carbon chain can occur at various points, producing fragments such as **m/z = 57** ($C_4H_9^+$, corresponding to the butyl group), **m/z = 43** ($C_3H_7^+$, corresponding to the propyl group), and **m/z = 29** ($C_2H_5^+$, corresponding to the ethyl group).

  Reactions:

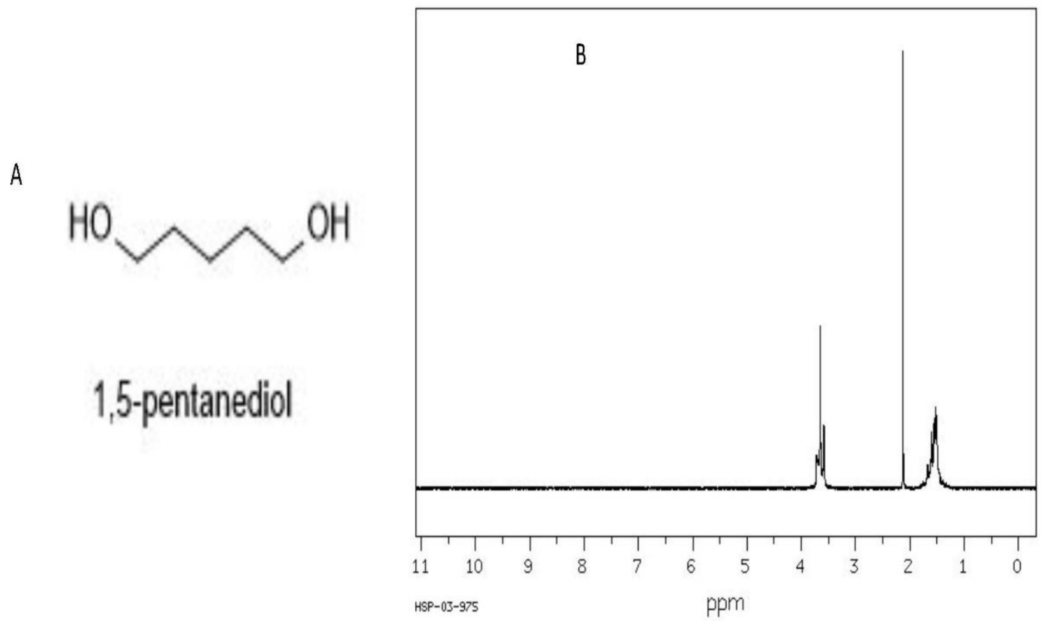
$$\mathrm{C_5H_{10}^+ \rightarrow C_4H_9^+ + CH_3} \quad (m/z = 57)$$
$$\mathrm{C_5H_{10}^+ \rightarrow C_3H_7^+ + C_2H_5} \quad (m/z = 43)$$
$$\mathrm{C_5H_{10}^+ \rightarrow C_2H_5^+ + C_3H_7} \quad (m/z = 29)$$

Summary of Expected Peaks:

- **m/z = 104 (M^+)**: Molecular ion.
- **m/z = 87**: Loss of one hydroxyl group (-OH).
- **m/z = 86**: Loss of water (H_2O).
- **m/z = 70**: Loss of both hydroxyl groups (-2OH).
- **m/z = 57**: Fragment corresponding to the butyl group ($C_4H_9^+$).
- **m/z = 43**: Fragment corresponding to the propyl group ($C_3H_7^+$).
- **m/z = 29**: Fragment corresponding to the ethyl group ($C_2H_5^+$).

These fragmentation patterns provide insight into the behavior of **1,5-pentanediol** during mass spectrometry, with the loss of hydroxyl groups and carbon chain cleavage being key processes leading to characteristic fragment peaks.

Figure S26: (A) Chemical structure of 1,5-pentanediol. (B) SDBS Database of 1,5-pentanediol (11). (C) Textual output from ChatGPT-4 for prominent fragments of 1,5-pentanediol spectra.



C

For **1,5-pentanediol**, which has two hydroxymethyl groups ($-CH_2OH$) at the 1- and 5-positions on a pentane chain, the ^1H NMR spectrum can be analyzed using the **n+1 rule** to predict the splitting patterns of the protons.

Structure Overview:

- **1,5-Pentanediol** has the following groups:
  - **Hydroxymethyl groups** ($-CH_2OH$) at positions 1 and 5
  - **Methylene groups** ($-CH_2-$) at positions 2, 3, and 4 in the middle of the chain

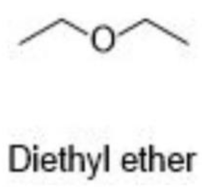
Proton Splitting Patterns Using the n+1 Rule:

1. **Hydroxymethyl Protons ($-CH_2OH$):**
   - The protons on the $-CH_2OH$ groups are adjacent to the central methylene groups but are not directly coupled with each other due to the distance. Each $-CH_2OH$ group is a distinct environment, and the -OH protons can potentially exchange and may not always be visible.
   - **Chemical shift:** Typically appear around **3.8 - 4.2 ppm**. If the -OH protons are observed, they would typically appear as a broad singlet around **2.0 - 4.0 ppm**.
2. **Methylene Protons ($-CH_2-$):**
   - For the methylene protons in the middle of the chain (positions 2, 3, and 4), we need to analyze the splitting pattern:
     - **Protons at positions 2 and 4** are each coupled to the adjacent protons (positions 3) but not to each other directly.
     - **Protons at position 3** are coupled to the adjacent methylene protons at positions 2 and 4.

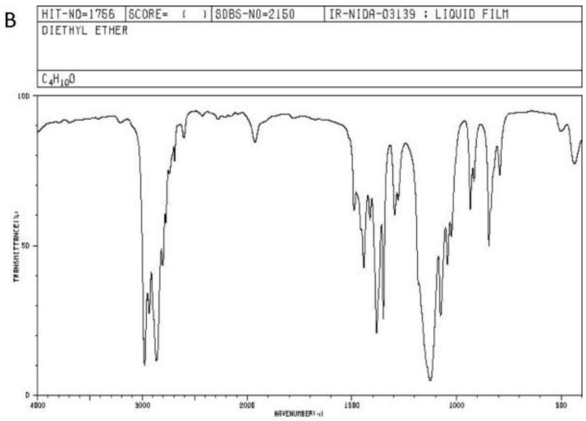
- **Position 2 and 4 Protons:**
  - Each of these methylene protons is coupled with two neighboring protons (one at position 3 on either side).
  - **Using n+1 Rule:** Each methylene proton at position 2 or 4 is split by 2 neighboring protons.
    - n = 2 (number of neighboring protons)
    - The result is a **triplet** for each of these protons.
  - **Chemical shift:** Typically appear around **1.6 - 2.0 ppm**.
- **Position 3 Protons:**
  - Each proton at position 3 is coupled with two neighboring protons (one at each side).
  - **Using n+1 Rule:** Each methylene proton at position 3 is split by 4 neighboring protons (2 on either side).
    - n = 4 (number of neighboring protons)
    - The result is a **quartet** for these protons.
  - **Chemical shift:** Typically appear around **1.6 - 2.0 ppm**, but can vary slightly due to their different chemical environment compared to positions 2 and 4.

Figure S27: (A) Chemical structure of 1,5-pentanediol. (B) SDBS Database of 1,5-pentanediol (11). (C) Textual output from ChatGPT-4 for prominent peaks of 1,5-pentanediol spectra.

A



B



C

For **diethyl ether ($C_4H_{10}O$)**, the IR spectrum will primarily reflect the characteristics of the **ether functional group** as well as the **aliphatic hydrocarbon chains**.

Structure of Diethyl Ether:

- Diethyl ether consists of two ethyl groups ($-C_2H_5$) connected by an oxygen atom, giving it the structure: **C_2H_5-O-C_2H_5**.

Key Peaks Expected in the IR Spectrum of Diethyl Ether:

1. **C-H Stretch (Alkyl, 2850–2960 cm^{-1}):**
   - The C-H stretching vibrations from the two ethyl groups will produce **strong peaks** in the **2850–2960 cm^{-1}** range.
   - This includes:
     - **Asymmetric stretching** around **2920–2960 cm^{-1}**.
     - **Symmetric stretching** around **2850–2900 cm^{-1}**.
2. **C-O Stretch (Ether, 1000–1200 cm^{-1}):**
   - The **C-O stretching vibration** from the ether functional group will produce a **strong peak** in the **1000–1200 cm^{-1}** range.
   - This peak is typically sharp and indicative of the presence of an ether.
3. **C-H Bending (Methylene, 1350–1470 cm^{-1}):**
   - The **bending (scissoring) vibrations** of the methylene ($-CH_2-$) groups will result in **moderate peaks** in the **1350–1470 cm^{-1}** range.
   - This is characteristic of aliphatic compounds.
4. **C-H Rocking (Methylene, 700–900 cm^{-1}):**
   - The **rocking motion** of the methylene ($-CH_2-$) groups gives rise to **moderate peaks** in the **700–900 cm^{-1}** range.
   - This peak indicates the presence of the longer aliphatic chains.

Summary of Key Peaks:

- **C-H Stretch (Alkyl)**: 2850–2960 cm^{-1} (strong).
- **C-O Stretch (Ether)**: 1000–1200 cm^{-1} (strong).
- **C-H Bending (Methylene)**: 1350–1470 cm^{-1} (moderate).
- **C-H Rocking (Methylene)**: 700–900 cm^{-1} (moderate).

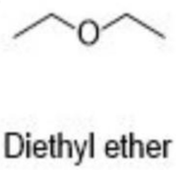
Additional Considerations:

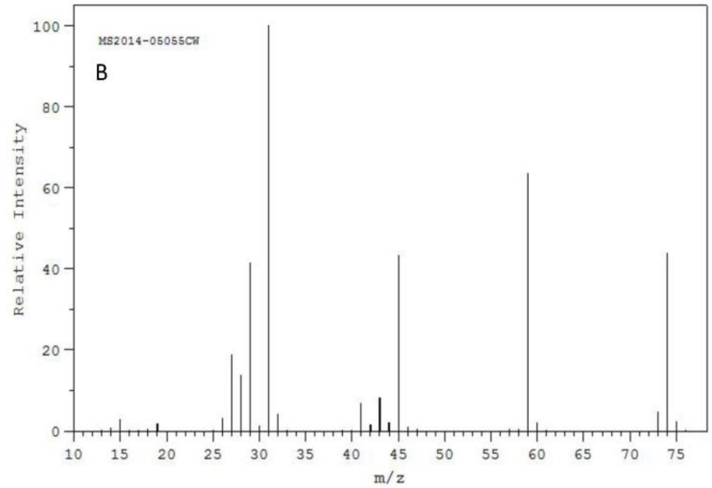
- **Ethers**: Diethyl ether is a simple ether, and its IR spectrum is typical of ethers, characterized by the strong C-O stretch and alkyl group features.
- **No O-H Peaks**: Unlike alcohols, diethyl ether will not show any O-H stretching peaks, making it distinct from alcohol-containing compounds.

This profile provides a comprehensive overview of the expected IR spectrum for diethyl ether. Let me know if you need any further details!

Figure S28: (A) Chemical structure of Diethyl ether. (B) SDBS Database of Diethyl ether (11). (C) Textual output from ChatGPT-4 for prominent peaks of Diethyl ether spectra.

A





C

Mass Spectrometry Analysis for Diethyl Ether ($C_4H_{10}O$)

Molecular Formula: $C_4H_{10}O$

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 4 = 48.04
- Hydrogen (H): 1.008 × 10 = 10.08
- Oxygen (O): 16.00 × 1 = 16.00

Total = 74.12 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M^+):

- The molecular ion for **diethyl ether** corresponds to the intact molecule and appears at **m/z = 74**. This peak represents the full diethyl ether molecule without fragmentation.

2. Loss of Ethyl Group (-C_2H_5):

- A common fragmentation pathway for diethyl ether involves the loss of one ethyl group (-C_2H_5, 29 Da), resulting in a fragment at **m/z = 45**. This fragment corresponds to the ethoxy ion ($C_2H_5O^+$).

  Reaction:

$$C_4H_{10}O^+ \rightarrow C_2H_5O^+ + C_2H_5 \quad (m/z = 45)$$

3. Loss of Methanol (-CH_3OH):

- Another possible fragmentation pathway is the loss of methanol (-CH_3OH, 32 Da), which results in a fragment at **m/z = 42**. This fragment corresponds to the remaining part of the molecule ($C_3H_6^+$).

  Reaction:

$$C_4H_{10}O^+ \rightarrow C_3H_6^+ + CH_3OH \quad (m/z = 42)$$

4. Loss of an Oxygen Atom (-O):

- Diethyl ether can also lose its oxygen atom during fragmentation, which results in a fragment at **m/z = 59**. This corresponds to the molecular ion without oxygen ($C_4H_{10}^+$).

  Reaction:

$$C_4H_{10}O^+ \rightarrow C_4H_{10}^+ + O \quad (m/z = 59)$$

5. Cleavage of the Ether Bond (-OCH_2):

- The ether bond (C-O-C) can break, resulting in the formation of two major fragments:
  - **m/z = 29**: Corresponding to the ethyl cation ($C_2H_5^+$).
  - **m/z = 45**: Corresponding to the ethoxy cation ($C_2H_5O^+$), as previously noted.

  Reaction:

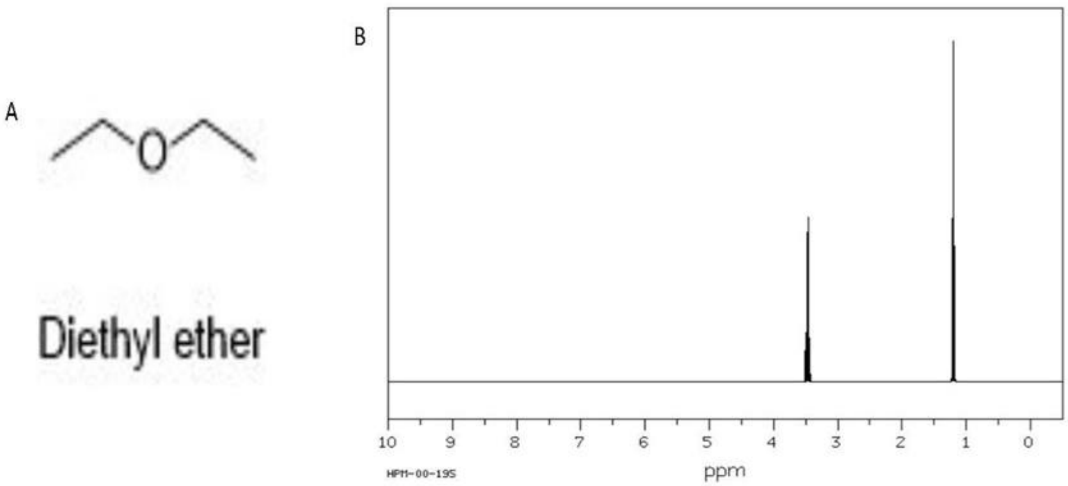
$$C_4H_{10}O^+ \rightarrow C_2H_5^+ + C_2H_5O^+ \quad (m/z = 29 \text{ and } 45)$$

Summary of Expected Peaks:

- **m/z = 74 (M^+)**: Molecular ion.
- **m/z = 45**: Loss of an ethyl group (-C_2H_5), resulting in the ethoxy ion ($C_2H_5O^+$).
- **m/z = 42**: Loss of methanol (-CH_3OH), producing the $C_3H_6^+$ fragment.
- **m/z = 59**: Loss of oxygen atom (-O).
- **m/z = 29**: Ethyl cation ($C_2H_5^+$) from cleavage of the ether bond.

These fragmentation patterns for **diethyl ether** provide clear insights into how the molecule breaks down under mass spectrometry conditions, with key peaks corresponding to the loss of ethyl groups, methanol, and oxygen, along with cleavage of the ether bond.

Figure S29: (A) Chemical structure of Diethyl ether. (B) SDBS Database of Diethyl ether (11). (C) Textual output from ChatGPT-4 for prominent fragmentation of Diethyl ether spectra.



C

Using the **n+1 rule** for **diethyl ether** ($C_2H_5OC_2H_5$), we can predict the splitting patterns of the proton signals:

Diethyl Ether Structure:

- **Ethyl group:** Each ethyl group has two types of protons:
  - **Methylene protons** (-CH_2-) next to the oxygen
  - **Methyl protons** (-CH_3) at the end of the ethyl chain

Proton Splitting Patterns:

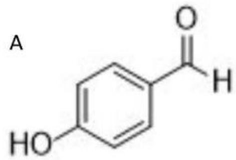
1. **Methylene Protons (-CH_2-):**
   - These protons are adjacent to the oxygen atom and couple with the three equivalent protons on the methyl group.
   - **Using n+1 Rule:** Each methylene proton is split by the three neighboring methyl protons
     - n = 3 (number of neighboring protons on the methyl group)
     - The result is a **triplet** (3+1=4) for the methylene protons.
   - **Chemical shift:** Typically around **3.4 - 3.7 ppm.**
2. **Methyl Protons (-CH_3):**
   - These protons are adjacent to the two equivalent methylene protons.
   - **Using n+1 Rule:** Each methyl proton is split by the two neighboring methylene protons.
     - n = 2 (number of neighboring protons on the methylene group)
     - The result is a **quartet** (2+1=3) for the methyl protons.
   - **Chemical shift:** Typically around **1.2 - 1.4 ppm.**

Summary Using the n+1 Rule:

- **Methylene protons (-CH_2-):**
  - **Triplet** at **3.4 - 3.7 ppm** (splitting due to 3 neighboring methyl protons).
- **Methyl protons (-CH_3):**
  - **Quartet** at **1.2 - 1.4 ppm** (splitting due to 2 neighboring methylene protons).

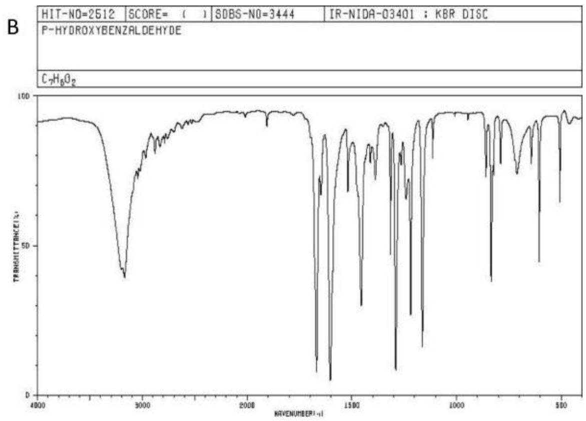
Figure S30: (A) Chemical structure of Diethyl ether. (B) SDBS Database of Diethyl ether (11). (C) Textual output from ChatGPT-4 for prominent peaks of Diethyl ether spectra.

A



4-hydroxybenzaldehyde

B



C

For **4-hydroxybenzaldehyde ($C_7H_6O_2$)**, the IR spectrum will exhibit features related to the **aldehyde functional group (-CHO)**, the **hydroxyl group (-OH)**, and the **aromatic ring**.

Structure of 4-Hydroxybenzaldehyde:

- **Aromatic ring**: A benzene ring with a hydroxyl group at the para position (4-position) and an aldehyde group at the 1-position.

Key Peaks Expected in the IR Spectrum of 4-Hydroxybenzaldehyde:

1. **C=O Stretch (Aldehyde, 1700–1750 cm^{-1})**:
   - The **carbonyl (C=O) stretch** from the aldehyde will produce a **strong, sharp peak** typically found in the **1700–1750 cm^{-1}** range.
   - This peak is characteristic of aldehydes and usually appears towards the lower end of this range due to interactions with the adjacent aromatic system.
2. **O-H Stretch (Hydroxyl, Broad, 3200–3600 cm^{-1})**:
   - The **hydroxyl (-OH)** group will produce a **broad, strong absorption** in the **3200–3600 cm^{-1}** region.
   - This broad peak results from hydrogen bonding and is typical for alcohols and phenolic compounds.
3. **C-H Stretch (Aromatic, 3000–3100 cm^{-1})**:
   - The aromatic C-H bonds will yield **weak to moderate peaks** in the **3000–3100 cm^{-1}** range, indicative of C-H stretching in aromatic systems.
4. **C-H Stretch (Aldehyde, 2700–2900 cm^{-1})**:
   - The aldehyde C-H stretch will produce a **weak to moderate peak** in the **2700–2900 cm^{-1}** range.
   - This corresponds to the C-H bond directly attached to the carbonyl carbon.
5. **C=C Stretch (Aromatic, 1450–1600 cm^{-1})**:
   - The **C=C stretching** vibrations in the aromatic ring will show **moderate peaks** in the **1450–1600 cm^{-1}** range.
   - These peaks arise from the conjugated double bonds present in the aromatic structure.
6. **C-H Bending (Aromatic, 700–900 cm^{-1})**:
   - The **out-of-plane bending** vibrations of the C-H bonds in the aromatic ring will produce **strong peaks** in the **700–900 cm^{-1}** range.
   - The specific pattern reflects the para substitution, often showing distinct peaks around **740–800 cm^{-1}**.
7. **C-O Stretch (Aromatic, 1200–1300 cm^{-1})**:
   - The **C-O stretching** vibration from the hydroxyl group may appear as a **weak to moderate peak** in the **1200–1300 cm^{-1}** range.

Summary of Key Peaks:

- **C=O Stretch (Aldehyde)**: 1700–1750 cm^{-1} (strong).
- **O-H Stretch (Hydroxyl)**: Broad, 3200–3600 cm^{-1} (strong).
- **C-H Stretch (Aromatic)**: 3000–3100 cm^{-1} (weak to moderate).
- **C-H Stretch (Aldehyde)**: 2700–2900 cm^{-1} (weak to moderate).
- **C=C Stretch (Aromatic)**: 1450–1600 cm^{-1} (moderate).
- **C-H Bending (Aromatic)**: 700–900 cm^{-1} (strong).
- **C-O Stretch (Aromatic)**: 1200–1300 cm^{-1} (weak to moderate).

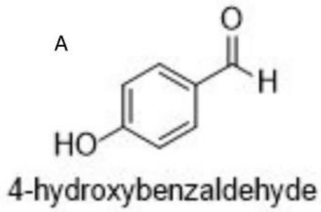
Additional Considerations:

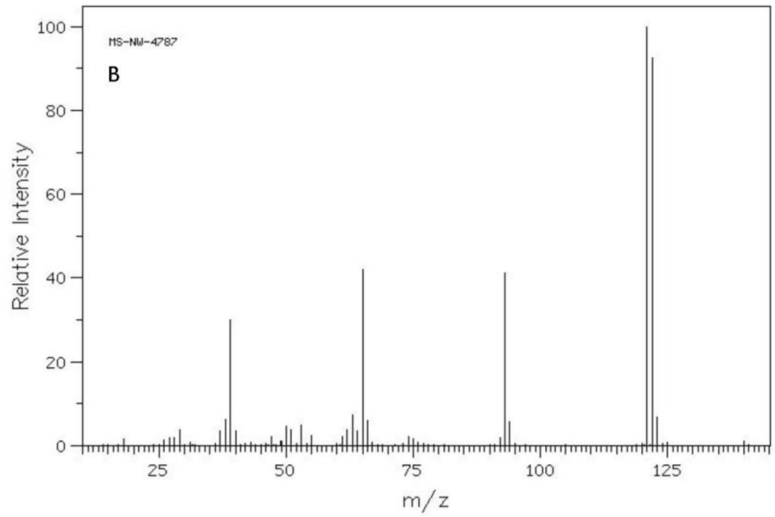
- **Hydrogen bonding**: The broad O-H peak indicates potential hydrogen bonding, which is common in phenolic compounds.
- **Aldehyde signature**: The distinct carbonyl peak confirms the presence of the aldehyde group, providing key identification for the compound.

This breakdown gives a detailed overview of the expected IR spectrum for 4-hydroxybenzaldehyde. Let me know if you need further information!

Figure S31: (A) Chemical structure of 4-hydroxybenzaldehyde. (B) SDBS Database of 4-hydroxybenzaldehyde (11). (C) Textual output from ChatGPT-4 for prominent peaks of 4-hydroxybenzaldehyde spectra.

A





C

Mass Spectrometry Analysis for 4-Hydroxybenzaldehyde ($C_7H_6O_2$)

Molecular Formula: $C_7H_6O_2$

Molecular Weight: Let's calculate the molecular weight:

- Carbon (C): 12.01 × 7 = 84.07
- Hydrogen (H): 1.008 × 6 = 6.048
- Oxygen (O): 16.00 × 2 = 32.00

Total = 122.12 g/mol

Key Peaks and Fragmentation Patterns

1. Molecular Ion (M⁺):

- The molecular ion for **4-hydroxybenzaldehyde** corresponds to the intact molecule and appears at **m/z = 122**. This peak represents the full structure of 4-hydroxybenzaldehyde.

2. Loss of Hydroxyl Group (-OH):

- A typical fragmentation pathway involves the loss of the hydroxyl group (-OH, 17 Da), resulting in a fragment at **m/z = 105**. This corresponds to the loss of the hydroxyl group, leaving the benzaldehyde portion of the molecule.

  Reaction:

$$C_7H_6O_2^+ \rightarrow C_7H_5O^+ + OH \quad (m/z = 105)$$

3. Loss of the Aldehyde Group (-CHO):

- Another common fragmentation involves the loss of the aldehyde group (-CHO, 29 Da), which results in a fragment at **m/z = 93**. This corresponds to the phenol ring without the aldehyde group.

  Reaction:

$$C_7H_6O_2^+ \rightarrow C_6H_5O^+ + CHO \quad (m/z = 93)$$

4. Cleavage of the Carbon-Carbon Bond Between the Aldehyde and Ring:

- The bond between the aldehyde group and the benzene ring can also cleave, yielding a fragment at **m/z = 77**, which corresponds to the phenyl cation ($C_6H_5^+$). This fragment is common in aromatic compounds.

  Reaction:

$$C_7H_6O_2^+ \rightarrow C_6H_5^+ + COH \quad (m/z = 77)$$

5. Loss of Carbon Monoxide (CO):

- The loss of a carbon monoxide (CO, 28 Da) fragment from the molecular ion can result in a fragment at **m/z = 94**, which corresponds to a structure where the aldehyde portion is partially fragmented, but the oxygen is still attached.

  Reaction:

$$C_7H_6O_2^+ \rightarrow C_6H_6O^+ + CO \quad (m/z = 94)$$

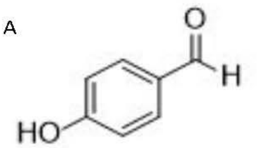
Summary of Expected Peaks:

- **m/z = 122 (M⁺)**: Molecular ion.
- **m/z = 105**: Loss of the hydroxyl group (-OH).
- **m/z = 93**: Loss of the aldehyde group (-CHO).
- **m/z = 77**: Phenyl cation ($C_6H_5^+$) from cleavage of the carbon-carbon bond between the aldehyde and the benzene ring.
- **m/z = 94**: Loss of carbon monoxide (CO).

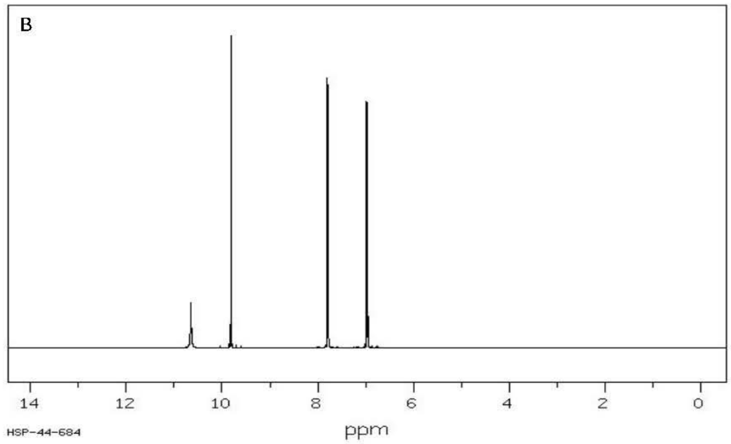
These fragmentation patterns for **4-hydroxybenzaldehyde** provide insights into the breakdown of the molecule, with key peaks related to the loss of functional groups like the hydroxyl and aldehyde, as well as the cleavage of the aromatic ring system.

Figure S32: (A) Chemical structure of 4-hydroxybenzaldehyde. (B) SDBS Database of 4-hydroxybenzaldehyde (11). (C) Textual output from ChatGPT-4 for prominent fragments of 4-hydroxybenzaldehyde spectra.

A



4-hydroxybenzaldehyde



C

Sure! Let's analyze the ^1H NMR spectrum for 4-hydroxybenzaldehyde using the n+1 rule.

Structure and Proton Environments

4-Hydroxybenzaldehyde ($C_7H_6O_2$) has the following proton environments:

1. **Aromatic protons (4 protons)**: These are the protons on the benzene ring.
2. **Aldehyde proton (1 proton)**: This is the proton on the -CHO group attached to the benzene ring.
3. **Hydroxyl proton (1 proton)**: This is the proton on the -OH group.

Analysis of Splitting Patterns

1. **Aromatic protons:**
   - The aromatic ring has a para-substituted pattern, which simplifies the splitting patterns.
   - For simplicity, the aromatic protons can be grouped as follows:
     - Two ortho protons relative to the hydroxyl substituent (split by 3 ortho/meta protons resulting in a doublet).
     - Two ortho protons relative to the aldehyde substituent (split by 3 ortho/meta protons, resulting in another doublet).
   - These protons typically appear in the range of δ 7.0-8.0 ppm.
2. **Aldehyde proton:**
   - The aldehyde proton is a singlet because it has no neighboring protons to split it. It typically appears downfield around δ 9.5-10.5 ppm.
3. **Hydroxyl proton:**
   - The hydroxyl proton often shows as a broad singlet due to hydrogen bonding and exchangeability with solvents. It typically appears around δ 4.5-5.5 ppm.

Simplified ^1H NMR Spectrum for 4-Hydroxybenzaldehyde

- **Aromatic region (δ 7.0-8.0 ppm)**: Two doublets due to the aromatic protons.
- **Aldehyde proton (δ 9.5-10.5 ppm)**: Singlet.
- **Hydroxyl proton (δ 4.5-5.5 ppm)**: Broad singlet.

Figure S33: (A) Chemical structure of 4-hydroxybenzaldehyde. (B) SDBS Database of 4-hydroxybenzaldehyde (11). (C) Textual output from ChatGPT-4 for prominent peaks of 4-hydroxybenzaldehyde spectra.