Proceedings of the 5th International Conference on Materials Chemistry and Environmental Engineering
DOI: 10.54254/2755-2721/2025.29832

An Overview of Retrosynthetic Analysis and Basic Guidelines

Qingyuan Ye

Ningbo Xiaoshi High School, Ningbo, China
imkellyye@gmail.com

Abstract. This paper provides an overview of basic retrosynthetic analysis, a powerful technique in organic synthesis developed by Elias James Corey in the late 1960s. Retrosynthetic analysis involves working backward from a target molecule, systematically breaking it down into simpler precursor molecules. By identifying the simplest reactions and the least number of steps needed to reassemble the target molecule, this approach enables more efficient synthesis of large, complex molecules. Guiding principles such as simplification, maximizing convergency, and exploiting symmetry are discussed, along with the consequences of bond breaking. Additionally, the study explores dioxygenation patterns, such as 1,2 and 1,3 dioxygenation patterns, etc., and talks about examples of retrosynthetic analysis for ibuprofen and paracetamol.

Keywords: retrosynthetic analysis, disconnection, dioxygenation pattern

1. Introduction

Elias James Corey, an American chemist, received the prestigious 1990 Nobel Prize in Chemistry for his groundbreaking work in the field of organic synthesis. His contributions to the theory and methodology of organic synthesis, particularly in the area of "retrosynthetic analysis," have had a profound impact on the field. This method streamlines the creation of complex, sizable molecules. In the late 1960s, Corey introduced a systematic approach, breaking away from the previous ad hoc methods.

In traditional synthesis, chemists started with simpler molecules and used trial and error to build up to the complex target molecule. Corey's retrosynthetic analysis, on the other hand, involves working backward from the target molecule, systematically breaking it down into simpler precursor molecules. This approach allows for more efficient synthesis by identifying the simplest reactions and the least number of steps needed to reassemble the target molecule [1].

Typically, retrosynthetic analysis is a method used to strategize the sequence of reactions needed to create a specific molecule. Working backwards from the known structure is an essential part of the process. The molecule to be synthesized, also known as the target molecule, is typically divided into two parts through a process called disconnection. The components are not molecules themselves, but rather fragments (referred to as synthons) that correspond to actual reagents [2].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.29832

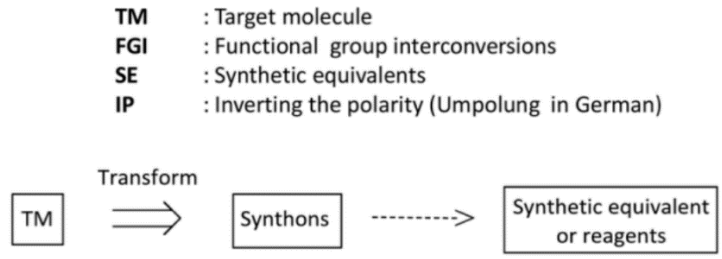


Figure 1. Basic retrosynthetic analysis

2. Background information

Retrosynthetic analysis is typically presented in the following manner, with the open arrows denoting the reverse direction of the analysis. Firstly, we identify precursor 1 that can react to obtain the target compound in a single step. Next, we determine precursor 2 that can react to produce the previous precursor, and we repeat this process until we have the starting material [3].

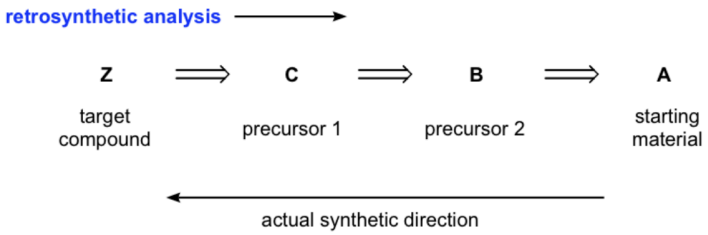


Figure 2. The direction of retrosynthetic analysis

The information on the "tree" (Figure 3) demonstrates the use of retrosynthetic analysis as a strategy for designing a synthesis in a reverse manner. Instead of starting with the starting material and determining the first step, we begin with the target molecule and determine the last step [3].

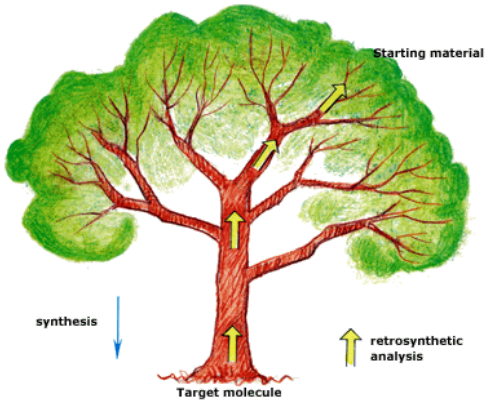


Figure 3. Retrosynthetic analysis is a "backward synthesis" [4]

DOI: 10.54254/2755-2721/2025.29832

3. Key concepts

3.1. Disconnection

Disconnection is a process that involves breaking bonds and converting a molecule into a potential starting material. It is essentially the opposite of a chemical reaction. In retrosynthesis, a specific arrow is used to indicate this process, as illustrated in Figure 4:

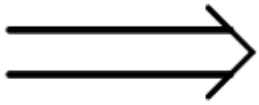


Figure 4. Retrosynthectic arrow

3.2. Target molecule

The complex organic molecule that one aims to synthesize. The process involves working backward from this target molecule to simpler precursor molecules, ultimately identifying the starting materials and the sequence of reactions needed to construct the target.

3.3. Synthon

A synthon is a general fragment, typically an ion, that is formed through a disconnection:

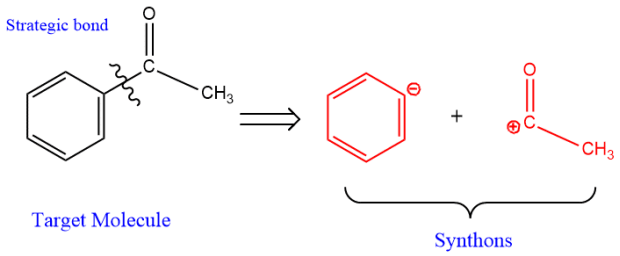


Figure 5. Synthons

3.4. Synthetic equivalent

Similar to a chemist's toolkit, a synthetic equivalent is a reagent that mimics the function of a synthon but has limitations due to its instability. It is essentially a combination of reagents that can carry out the desired transformation represented by a synthon.

DOI: 10.54254/2755-2721/2025.29832

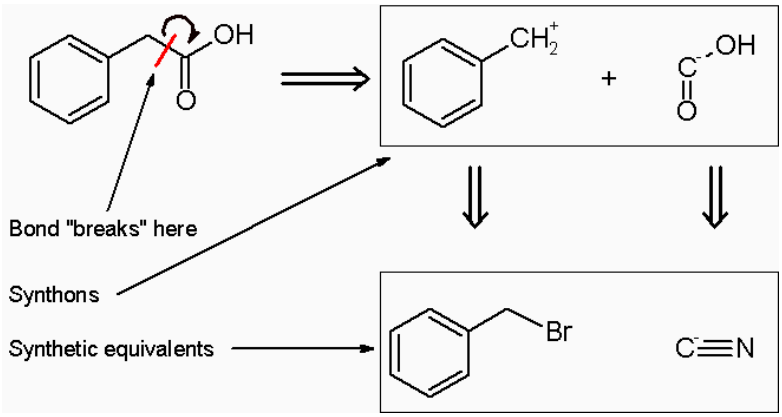


Figure 6. Difference between synthons and synthetic equivalents

4. Guiding principles

4.1. Simplification

For optimal results, it is crucial to simplify the complex target molecule at each step of the retrosynthetic process, making it resemble the starting material as closely as possible.

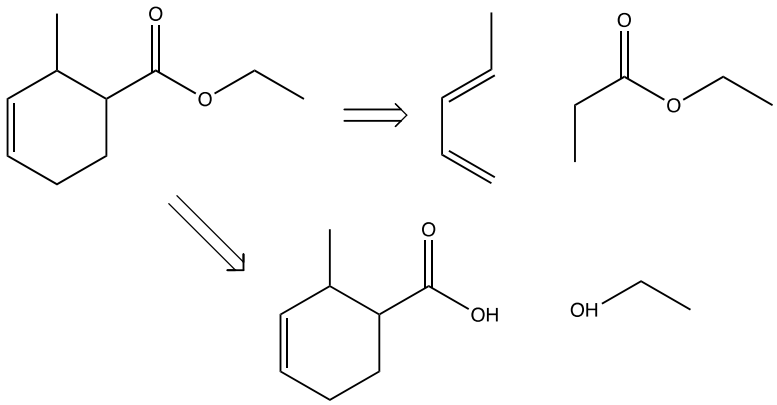


Figure 7. Simplification by breaking bonds

4.2. Maximize convergency

When it comes to designing synthesis, the convergent pathway is always the preferred choice over the linear pathway. This is because it helps save steps and simplifies the overall process. Additional procedures can decrease the overall quantity of the initial substances as each step's chemical efficiency impacts the ultimate yield, which is determined using power functions. Just like a chemist, the convergent pathway serves the purpose of breaking down complex molecules into similar fragments. Using the compound "A-B-C-D-E-F," we can demonstrate the mechanisms of both convergent and linear synthesis processes. When utilizing a linear synthesis route, a total of five steps are necessary to achieve the target molecule. However, in convergent synthesis, the approach is designed to split in the middle, resulting in a reduction in the overall number of steps required, down to just three. Efficient and effective results can be obtained by utilizing convergent synthesis [5].

DOI: 10.54254/2755-2721/2025.29832

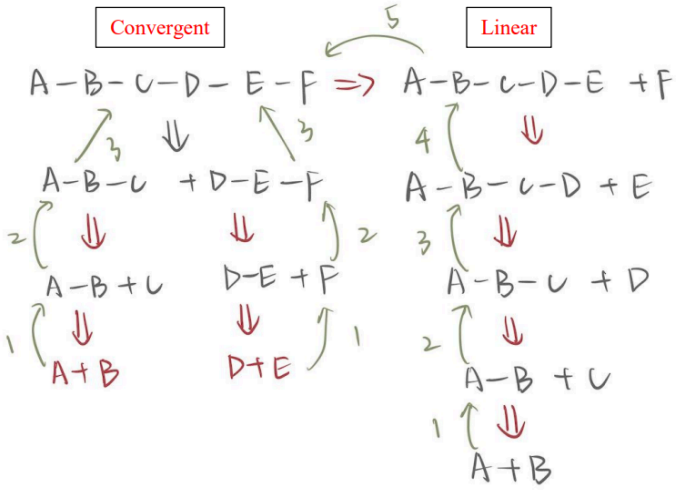


Figure 8. Convergent synthesis on the left and linear synthesis on the right

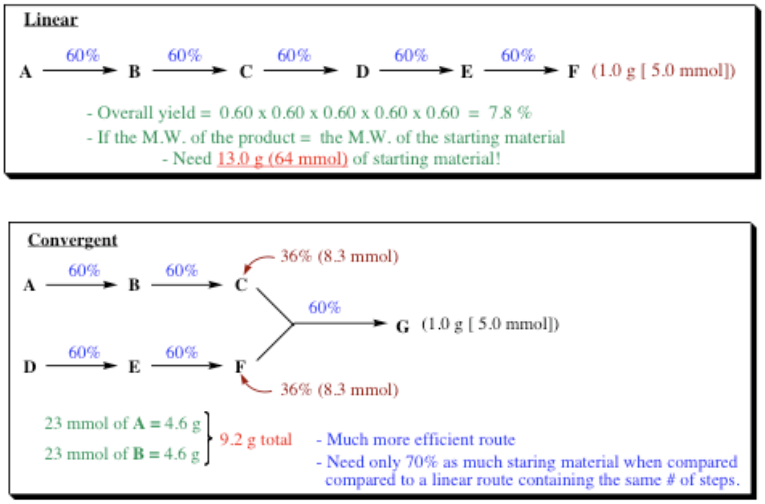


Figure 9. Comparison with linear synthesis and convergent synthesis

4.3. Fine tuning

4.3.1. Functional Group Interconversion (FGI)

Functional group interconversion (FGI) is the process of transforming one functional group into another through various chemical reactions, such as substitution, addition, elimination, oxidation, or reduction. This process is also utilized in retrosynthetic analysis, where the reverse reactions are considered [5].

DOI: 10.54254/2755-2721/2025.29832

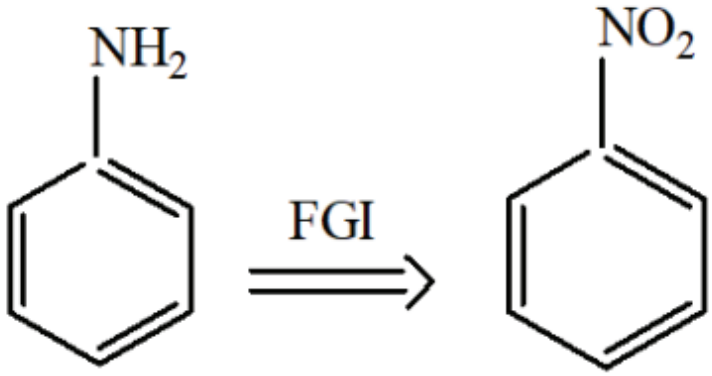


Figure 10. Functional group interconversion

For example, FGI is used in the retrosynthetic analysis for ofornine shown in Figure 11 [2]:

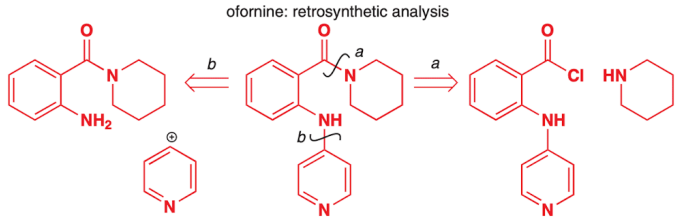


Figure 11. Retrosynthetic analysis for ofornine [2]

Functional group interconversion (FGI) can be a useful tool for making disconnections in a compound. In some cases, it may be preferable to avoid disconnecting reactive functional groups like acyl chlorides or alkyl halides, as they can cause issues with chemoselectivity. In addition, they can be quite handy when dealing with molecules that have challenging functional groups to disconnect [2].

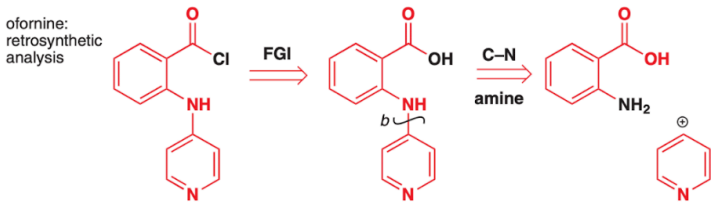


Figure 12. Retrosynthetic analysis for ofornine [2]

4.3.2. Functional Group Addition (FGA)

It introduces a functional group to the original compound. Certain functional groups require the addition of a group to the immediate precursor that can be easily disconnected. For instance, the dehydration of a β-hydroxy ketone results in the formation of an α, β-unsaturated ketone [6].

DOI: 10.54254/2755-2721/2025.29832

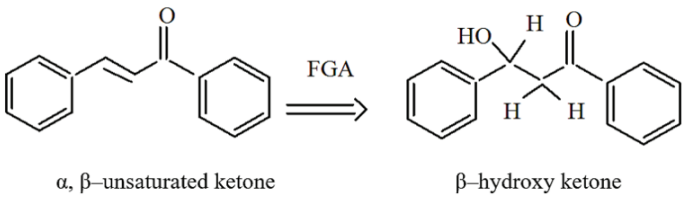


Figure 13. Functional group addition

4.3.3. Functional Group Removal (FGR)

It removes a functional group from the original compound. In the following example, the halide group is removed to obtain a precursor that contains one carbonyl functional group.

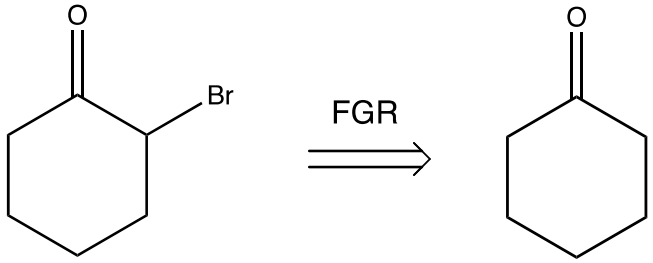


Figure 14. Functional group removal

4.3.4. Two-group disconnections are better than one-group disconnections

In Figure 15, the alcohol is adjacent to the ether we wanted to disconnect, and we included both of them in the disconnection [3]. Using a two-group disconnection approach allows for the simultaneous activation of both functional groups, resulting in a more efficient return to simple starting materials [2]. It has the potential to streamline the process and boost the final yield.

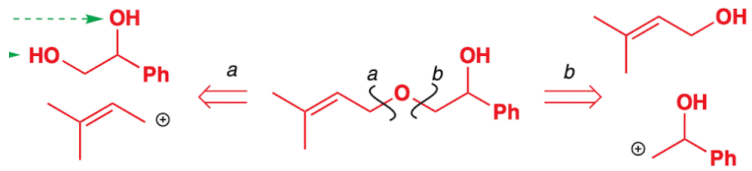


Figure 15. Two-group disconnections

4.4. Exploit symmetry when possible

Understanding the symmetrical properties of a molecule can greatly simplify the retrosynthetic process and reduce its overall complexity. This point is connected to the previous point about maximizing convergency, as symmetry presents a chance for a convergent strategy [7].

DOI: 10.54254/2755-2721/2025.29832

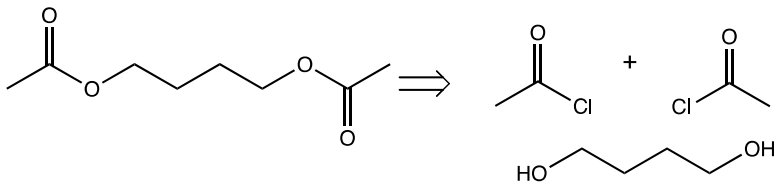


Figure 16. Symmetry within a molecule

5. Consequences of bond breaking

When two electrons are shared in a single bond, their disconnection leads to a variety of outcomes. Whenever a single bond is broken, there are three potential scenarios for electron movement, and each of these outcomes has a direct impact on the synthon that is produced.

The first case is that the left synthon is positively charged, and the right synthon is negatively charged (Figure 17).

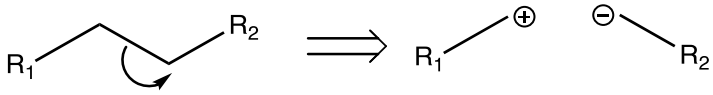


Figure 17. Positively-charged synthon on the left and negatively-charged on the right

The second case is that the right synthon is negatively charged, and the left synthon is positively charged (Figure 18), resulting from the opposite moving direction of electrons.

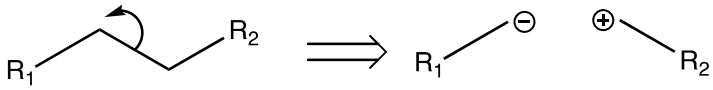


Figure 18. Positively-charged synthon on the right and negatively-charged on the left

Another case involves free radicals (Figure 19). Free radicals are typically highly reactive as they eagerly seek an additional electron to occupy their orbital.

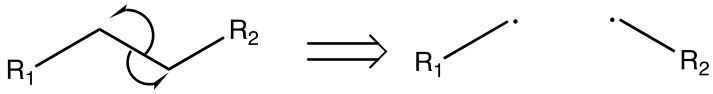


Figure 19. Free radicals are formed

6. Dioxygenation patterns

6.1. 1,2 dioxygenation pattern

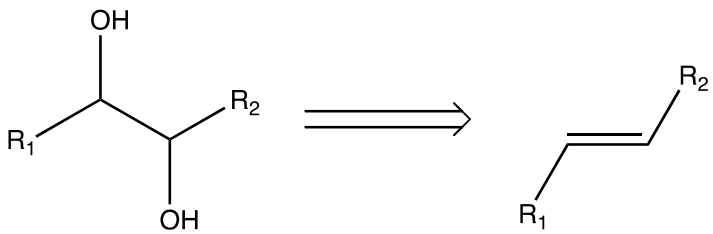


Figure 20. 1,2 dioxygenation pattern

DOI: 10.54254/2755-2721/2025.29832

When dealing with target molecules that exhibit the 1,2 dioxygenation pattern, the initial step involves transforming the molecule into its canonical structure of 1,2-diol through functional group interconversion (FGI). Afterwards, the diol is disconnected to create the alkene. This reaction occurs when it interacts with osmium tetroxide (OsO_4), resulting in the formation of diolate species that can be hydrolyzed to produce the cis-diol. The process of converting the alkene to the diol in the presence of OsO_4 is known as dihydroxylation [8].

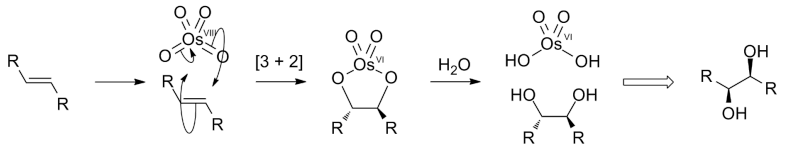


Figure 21. Mechanism of dihydroxylation

In the synthesis of 1,2-diol, the Pinacol coupling reaction is introduced, which exhibits poor selectivity because the reaction proceeds through radical intermediates that can combine in different orientations, leading to a mixture of syn and anti-diols.

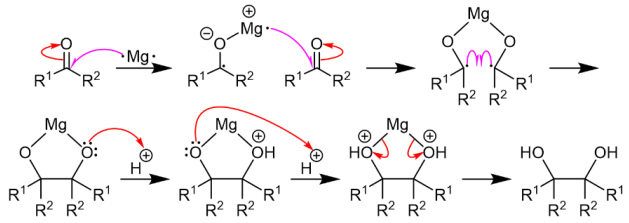


Figure 22. Mechanism of pinacol coupling

And the 1,2-disconnection of diol (pinacol rearrangement) is shown in Figure 23, involving the conversion of a 1,2-diol (pinacol) to an aldehyde or ketone:

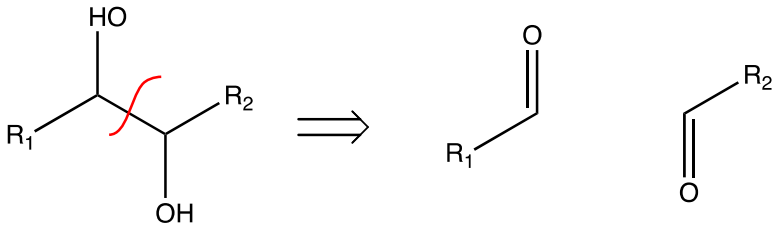


Figure 23. Disconnection of 1,2-diol

6.2. 1,3 dioxygenation pattern

Hydroxyketones, or α,β-unsaturated ketones that can be interconverted with hydroxyketones, are canonical structures of molecules that follow the 1,3 dioxygenation pattern. After the target molecule undergoes functional group interconversion (FGI), the 2,3 bonds of the hydroxyketone are disconnected to generate ketones and aldehydes.

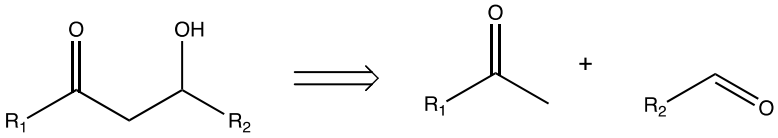


Figure 24. 1,3 dioxygenation pattern

DOI: 10.54254/2755-2721/2025.29832

In order to generate the canonical structure, the carbonyl group must undergo base-catalyzed reaction to form an enolate. Just as a chemist would expect, the enolate undergoes a reaction with the carbonyl group known as an aldol addition. If an aldol addition takes place, the resulting product is a hydroxy ketone. However, if an aldol condensation occurs in addition to the addition, the product will be an α-β unsaturated ketone.

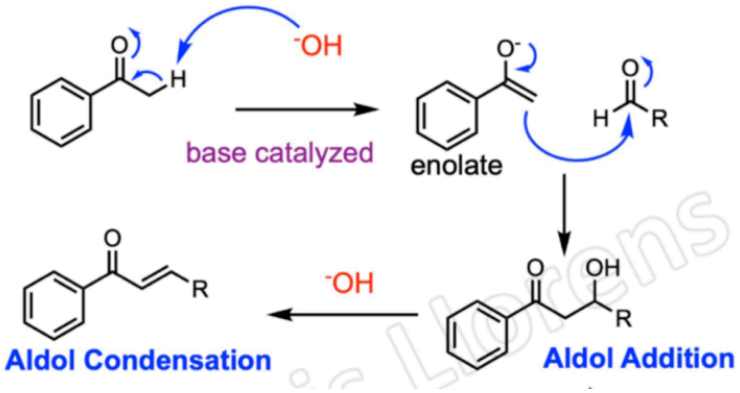


Figure 25. Mechanism of aldol condensation [9]

6.3. 1,4 dioxygenation pattern

The canonical structure of the 1,4 dioxygenation pattern is 1,4-dicarbonyl. The target molecule is firstly transformed into this canonical structure. The structure undergoes disconnections by breaking the 2,3 bond. Through a series of reactions, a disconnection occurs resulting in the formation of two parts: a methyl carbonyl and a halomethyl ketone. These parts can then form the canonical structure [9].

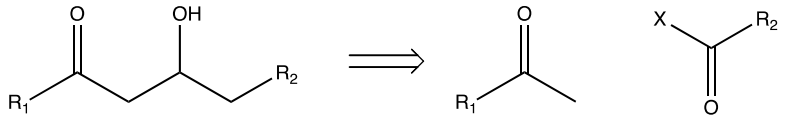


Figure 26. 1,4 dioxygenation pattern

Begin by treating the methyl ketone with a base to generate an enolate known as enolization. In another reaction, the enolate combines with a halomethyl ketone to produce the 1,4-dicarbonyl compound.

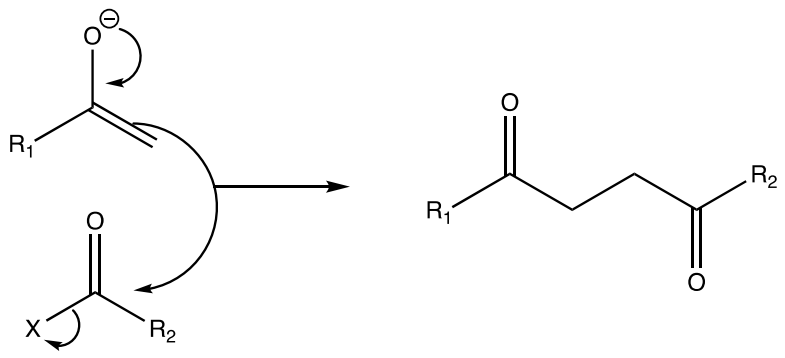


Figure 27. Mechanism of enolization

DOI: 10.54254/2755-2721/2025.29832

6.4. 1,5 dioxygenation pattern

A 1,5-dicarbonyl compound follows the 1,5-dioxygenation pattern. It is electronically consonant, so the electronic distribution within the molecule is stabilized. The first step is to perform functional group interconversion (FGI) on the target molecule. Subsequently, the 2,3 bond or 3,4 bond in the 1,5-dicarbonyl compound is disconnected, resulting in a methyl carbonyl and an α-β unsaturated system.

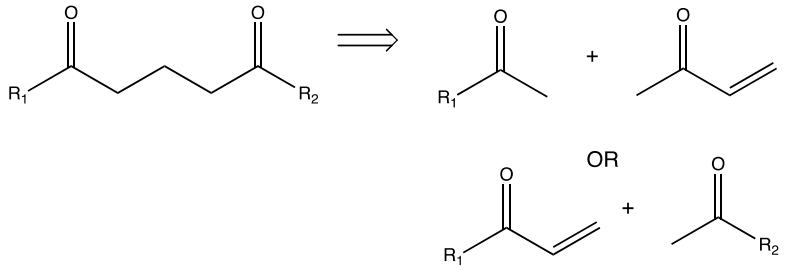


Figure 28. 1,5 dioxygenation pattern

In the synthesis of 1,5-dicarbonyl compounds, the methyl carbonyl is first enolized to form an enolate. The enolate then reacts with an α-β unsaturated system via Michael addition (part of Robinson annulation) to form the 1,5-dicarbonyl compound (the target molecule).

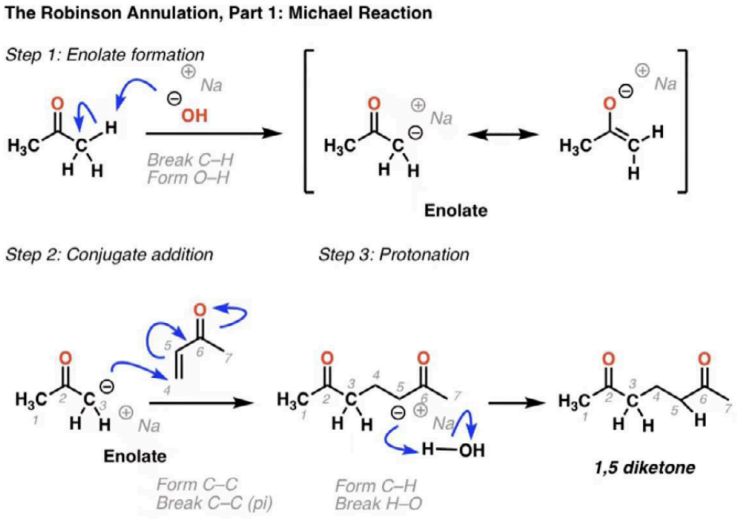


Figure 29. Mechanism of michael addition [7]

6.5. 1,6 dioxygenation pattern

1,6 dioxygenation is an illogical disconnection where two carbonyl groups react to create cyclohexene. To begin, we must transform the target molecule into a 1,6-dicarbonyl structure through functional group interconversion (FGI). This particular structure is considered a canonical structure that follows a 1,6 dioxygenation pattern.

DOI: 10.54254/2755-2721/2025.29832

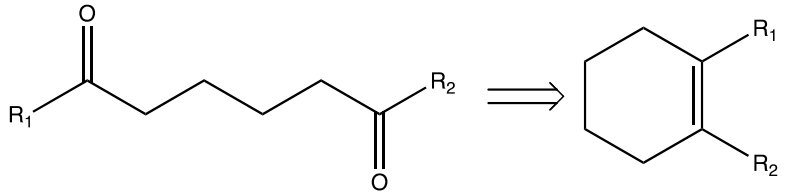


Figure 30. 1,6 dioxygenation pattern

The reverse reaction can be completed by ozonolysis (cyclohexene reacts with O_3):

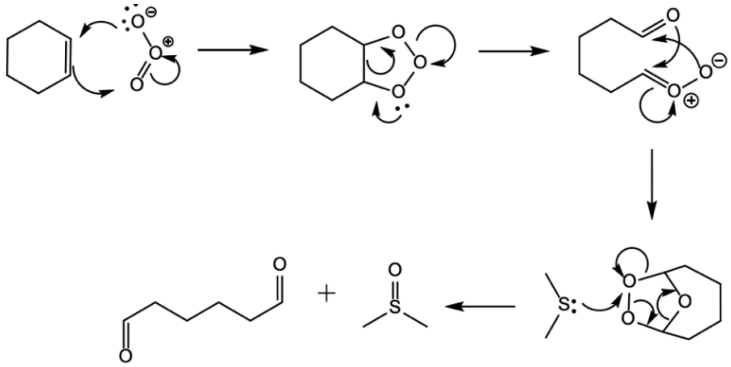


Figure 31. Mechanism of ozonolysis

7. Other examples of retrosynthetic analysis

7.1. Retrosynthetic analysis of ibuprofen

The designed retrosynthetic analysis for ibuprofen involves functional group interconversion (FGI). Sometimes, a disconnection carried out does not generate sufficiently stable fragments, but such fragments can be obtained by FGI. The repeated C-C disconnections are used to obtain synthons, which help construct the synthetic equivalents and, thus, the starting materials for retrosynthesis. The retrosynthesis route can involve Fridel Crafts acylation and Clemenson reduction.

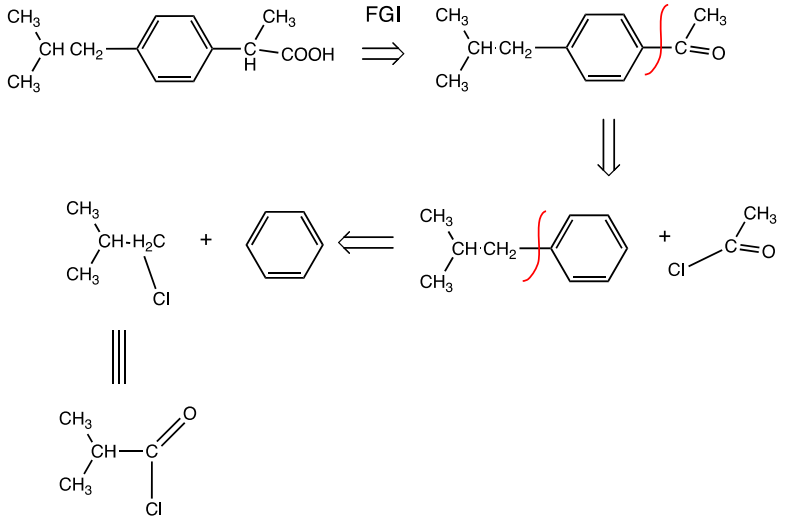


Figure 32. The retrosynthetic analysis for ibuprofen

DOI: 10.54254/2755-2721/2025.29832

7.2. Retrosynthetic analysis of paracetamol

As shown in Figure 33, the target molecule A can be broken down into synthons B and C. Two synthetic molecules, D and E, have the ability to undergo a straightforward nucleophilic addition-elimination reaction when reacted together. And then molecule D can be further disconnected into synthons F and G. NH_2 is a poor electrophile (it usually acts as a nucleophile). Thus, a synthetic equivalent electrophile for NH_2 could be NO_2, which has the ability to be reduced to NH_2. Thus NO_2 can react with phenol (molecule H) through electrophilic substitution [10].

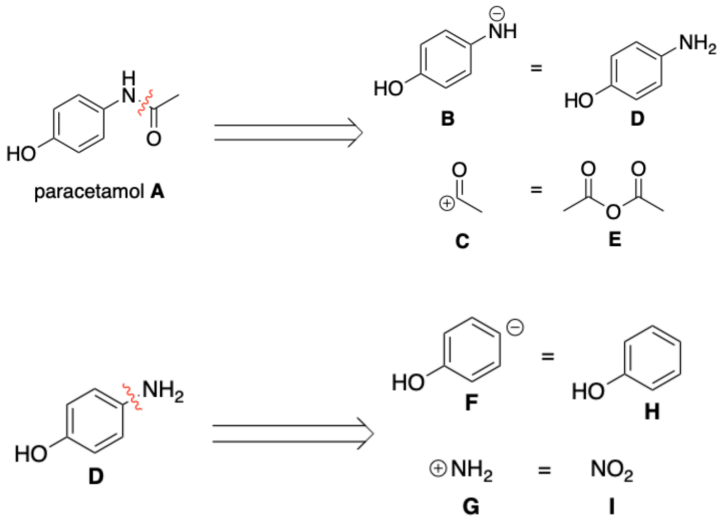


Figure 33. The retrosynthetic analysis for paracetamol [10]

8. Conclusion

Overall, retrosynthetic analysis is a valuable tool in the field of organic synthesis, revolutionizing the approach to complex molecule synthesis. This systematic and strategic method allows chemists to work backward from a target molecule, breaking it down into simpler precursor fragments. The concept of disconnection, where bonds are broken to convert a molecule into possible starting materials, lies at the core of this approach. By identifying the highly efficient reactions and the fewer steps required, it enables the synthesis of large and intricate molecules in a more streamlined and cost-effective manner.

References

[1] Shampo, M. A., Kyle, R. A., & Steensma, D. P. (2013). Elias James Corey—Nobel Prize for Retrosynthetic Analysis. Mayo Clinic Proceedings, 88(1), e7.
[2] Clayden, J., Greeves, N., & Warren, S. (2012). Organic Chemistry (2nd ed., Vols. 694–722). Oxford University.
[3] Liu, X. (n.d.). 9.6: Synthesis of Target Molecules- Introduction of Retrosynthetic Analysis. Libretexts.
[4] Nobel Prize Outreach AB 2024. (n.d.). The Nobel Prize in Chemistry 1990. NobelPrize.org. Retrieved September 20, 2024, from https: //www.nobelprize.org/prizes/chemistry/1990/8823-retrosynthetic-tree/
[5] Warren, S. (1991). Designing organic syntheses: A Programmed Introduction to the Synthon Approach. John Wiley & Sons.
[6] Badami, B. V. (2019). Retrosynthetic analysis. Resonance : Journal of Science and Education, 24(10), 1071–1086. https: //doi.org/10.1007/s12045-019-0877-2
[7] Ashenhurst, J. (2023, June 16). The Michael addition reaction and conjugate addition. Master Organic Chemistry. https: //www.masterorganicchemistry.com/2023/05/24/michael-addition-reaction-conjugate-addition/

DOI: 10.54254/2755-2721/2025.29832

[8] Liu, H., & Zhu, J. (n.d.). Introduction of Retrosynthetic Analysis and Guidelines Application of Analyze Molecules. Dean & Francis.
[9] Llorens, L. (2024, January 9). Aldol reaction - nrochemistry.com. nrochemistry.com - Named Reactions in Organic Chemistry. https: //nrochemistry.com/aldol-reaction/
[10] Retrosynthesis (Aspirin and paracetamol). (2020, June 22). WordPress.com. https: //chembam.com/resources-for-students/the-chemistry-of/retrosynthesis-aspirin-paracetamol/