Basics of Retrosynthetic Analysis: A Review for Beginners

Jingyuan Wang, Yifan Ji, Yunzheng Tian, Tianqi Fang[4]

Abstract

Retrosynthesis in organic chemistry seeks a possible way to synthesize a complex organic molecule by considering the reverse synthesis route. Chemists can explore multiple possibilities and find the most efficient synthesis route by starting from the target molecule and returning to simple reactants. This work is an essential guide for doing retrosynthesis analysis by providing examples and explanations of the fundamental notions and procedures involved. We will discuss topics like bond disconnections, FGIs, and synthons. This work illustrates the retrosynthesis analysis of complicated organic molecules using the discussed principles and methods.

Keywords: Retrosynthesis, Synthon, FGI, Bond Disconnection

1. Introduction

Organic synthesis is a crucial part of chemistry that can directly lead to scientific and societal benefits by actualizing new molecules that have ideal functions in different disciplines, such as material science and medicine. Retrosynthetic analysis is formalized in E.J. Corey's book The Logic of Chemical Synthesis [1]. Retrosynthesis can be applied in actual laboratory practices as a technique that can be used to break down complex molecules and help synthesize organic molecules. Chemists widely use this method to increase efficiency and feasibility when synthesizing new molecules while making the process more environmentally friendly. These advantages are a basis for accelerating drug discovery and will have long-lasting impacts in multiple fields of study [2]. Therefore, this work below would conclude the main idea of retrosynthesis by emphasizing certain approaches and critical terms, aiming to offer an indication to beginners with basic theoretical knowledge.

2. What is retrosynthesis

Retrosynthetic analysis is a method and an intellectual tool adopted to help come up with a synthesis for a complex organic molecule. The retrosynthetic analysis allows chemists to conduct the "reverse" synthesis process. Starting from the target molecule, it "deconstructs" the large and complicated target molecule into many simpler and more basic constituents, known as synthons. Based on known reactions, retrosynthesis generates a roadmap that guides the synthesis. Retrosynthesis also helps explore and compare the different possible routes to synthesize the target molecule, which helps find the most suitable synthesis method.

A general retrosynthetic analysis is shown in Figure 1.

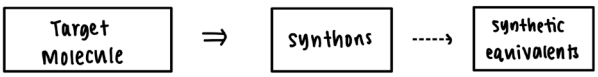

Figure 1. A basic retrosynthetic analysis

2.1 *Benefits of doing retrosynthesis:*

More economical. Conducting retrosynthetic analysis before doing the synthesis straightaway allows different routes to be compared based on feasibility and the cost of related reagents. Thus, the more efficient and effective pathway, particularly cost-effective, can be used for large-scale synthesis.

2.2 General principles when conducting retrosynthesis:

2.2.1 Focus on simplification

The primary purpose of retrosynthesis is to develop a synthesis route from basic molecules, so the retrosynthesis route that ends up with more simple molecules will be considered a better method.

2.2.2 Maximize convergency

Retrosynthesis can be grouped into two major strategies——linear and convergent.

In a linear strategy, the target molecule is retrosynthesized through a series of linear transformations. In a convergent approach, the target molecule is split into two similar halves that are retrosynthesized independently. A convergent strategy is preferred as there will be fewer steps in total and leads to a higher overall yield, meaning it is more efficient.

In Figure 2, the convergent method only takes three linear steps from molecule A to the target molecule, while the linear strategy takes five linear sequences.

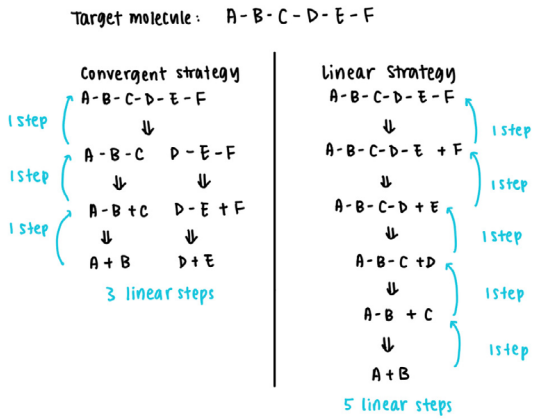

Figure 2 Comparison of the retrosynthetic process using the convergent strategy with linear strategy

2.2.3 *Use a good and reasonable mechanism*

Good retrosynthetic routes contain likely intermediates in the reaction, making the synthesis feasible.

For instance, in Figure 3, there are two possible retrosynthetic pathways for the target molecule below. However, route one is preferred as Me+ and Me2COH- involved in the second route are unlikely species in reality.

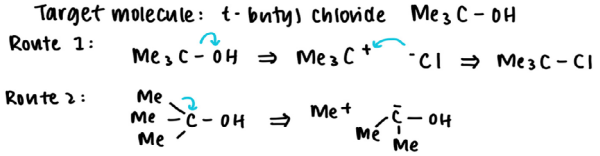

Figure 3 Retrosynthetic analysis of t-butyl chloride (Me_3C-OH)

2.2.4 *Two group disconnections are better than one group disconnections*

As two group disconnections can lead to more simplification, these disconnections are preferred over only one group disconnections.

2.2.5 *Minimize the use of fine-tuning*

Protecting groups (such as forming a cyclic ketal) for fine-tuning adds to the synthetic process. It decreases the simplicity of retrosynthesis, so use them only if necessary.

2.2.6 *Exploit symmetry*

Exploiting symmetry within a molecule significantly reduces its complexity and simplifies the retrosynthetic process. This point relates to the previous point of maximizing convergency as symmetry introduces an opportunity for a convergent strategy.

As shown in Figure 4, this molecule can be best disconnected by separating the symmetrical carboxylic group on the ends.

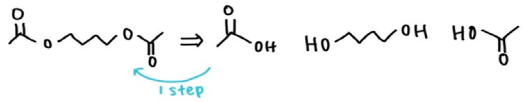

Figure 4 Bond disconnection of the target molecule.

2.3 Symbols of retrosynthesis:

Retrosynthetic arrow: Represents going backward from the target molecule to simpler molecules. It is shown in Figure 5.

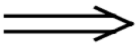

Figure 5 Retrosynthetic arrow

Synthetic arrow: Represents going forwards from the simpler molecules to the target molecule. It is shown in Figure 6.

Figure 6 Synthetic arrow

Bond disconnection: A wavy line ( ) to show where the bond is being broken (disconnected)

The wavy lines beside oxygen mark the bond disconnections in Figure 7.

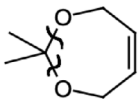

Figure 7 Bond disconnection is indicated by the wavy lines

In retrosynthesis, the target molecule is broken down through bond disconnections and functional group interconversions (FGI), illustrated below.

3. Bond disconnections

3.1 Common atom approach

As mentioned above, while considering bond disconnection, simplifying the complicated structure of the molecule is crucial. In this case, a strategic device called the 'common atom approach,' used in polycyclic compounds, could be used as a technique for people to reduce complexity. It's summarized as when breaking bonds in a polycyclic molecule, the one between two common atoms is supposed to be considered first to make the most strategic disconnections. The 'common atoms' are the atoms involved in more than one ring.

3.2 One-group disconnections

One-bond disconnection is a bond disconnection that can split molecules with a relatively simple structure. The common types are the disconnection of alcohols, olefins,

acids, and some carbon compounds. For instance, C-C disconnections are one form. The C-C bond adjacent to an alcohol group can be disconnected, ending with an aldehyde and a Grignard reagent [3].

Figure 8 shows a general disconnection of C-C in tertiary alcohol [4].

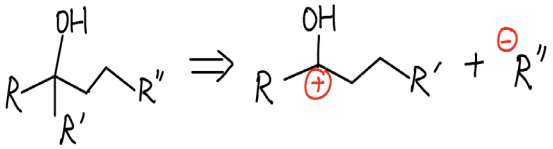

Figure 8 the retrosynthetic pathway of tertiary alcohol

3.3 Deoxygenation: two-group disconnections

Apart from one-bond disconnection, the other imaginary bond disconnection method is two-bond disconnection. It can be used when target molecules have two functional groups. And it works better than the last method.

One of the common techniques within two-bond disconnection is called deoxygenation, and it contains the following patterns.

(1) 1,2-deoxygenation

This approach can be applied when the functional groups are connected with carbon one and carbon 2 in the chain of the target molecule. The bond between the functional group and the oxygenated carbon could be a single bond or a double bond. The general structure of the TM might be different, so the first necessary step is converting the molecule back into its prototypical version by functional group interconversion, which is a diol, and the end product would be an olefin. As figure 9 shows, one of the products is an olefin [3]. This retrosynthetic route is based on the forward reactions, which are hydroxylation of the olefin, and for which reactions, reagents such as OSO4 are mainly used.

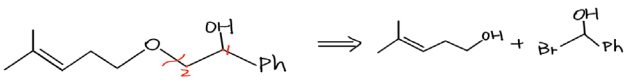

Figure 9 retrosynthetic route based on the forward reaction of hydroxylation and olefin

(2) 1,3-deoxygenation

For this pattern, oxygen is placed on carbon1 and carbon 3. The first step would be FGI to convert the molecule back into a canonical version of a hydroxy ketone if needed. The outcome of this retrosynthetic process is generally a ketone and an aldehyde.

(3) 1,4/ 1,5/ 1,6-deoxygenation

The oxygen in these cases is located at 1,4 carbon, 1,5 carbon, and 1,6 carbon, respectively, and the canonical molecule of 1,4 1,5 1,6-deoxygenation is all dicarbonyl pieces. The forward reactions of 1,4 1,5 1,6-deoxygenation are alkylation reaction, Michael addition, and ozonolysis. There are multiple ways to disconnect the bonds, but the outcome of these routes are similar.

An example of 1,6-deoxygenation is shown in figure 10 [5].

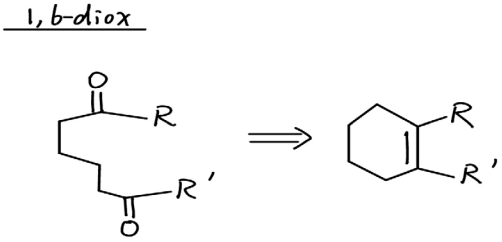

Figure 10. 1,6 dioxygenation

These patterns introduce some classical ways people would choose during retrosynthesis to make the TM.

4. FGI

4.1 ***What is FGI:***

FGI stands for functional group interconversions. It describes the process of changing one functional group to another functional group. In a retrosynthetic analysis, FGI forms the foundation of later bond disconnections.

4.2 ***Common FGIs:***

Disconnections involving H- are mostly redox reactions instead of disconnections. They are examples of FGIs. Common FGIs are shown in Figure 11 and Figure 12. An example of retrosynthetic analysis using FGI is shown in Figure 13.

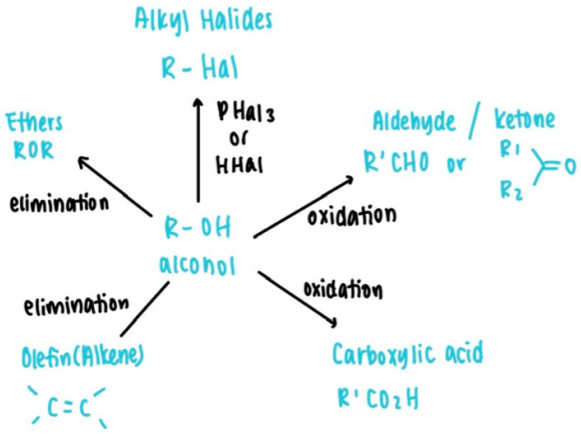

Figure 11 Common FGIs related to alcohol

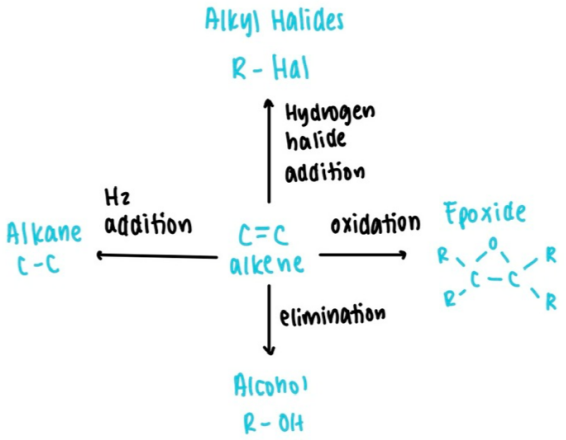

Figure 12 Common FGIs related to olefin

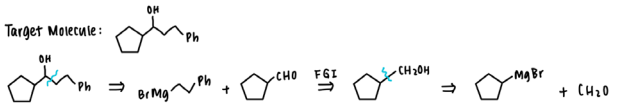

Figure 13 Example of using FGI in retrosynthesis

5. Synthon

5.1 Synthon definition:

A synthon is a constituent part of the target molecule that can be synthesized using starting materials. It acts as the "building block" [6]. Synthons can either be one tiny atom or a complex molecule nearly the size of the target molecule. In addition, atoms in the target molecule can be shared by multiple synthons. Those that are shared by several rings are called "common atoms."

Different synthetic ways can be developed depending on the synthons recognized; therefore, the complexity of the synthesis is associated with the numbers, sizes, shapes, and types of synthons identified. Furthermore, certain synthons within a target molecule are advantageous because of the ease of obtaining these molecules (often commercially or laboratory-wise favorable). Therefore, choosing the inexpensive and available synthons is crucial while intending to simplify the target molecule. For example, it would simplify the synthetic process for symmetric molecules to disconnect the symmetrical parts, as shown in figure 14.

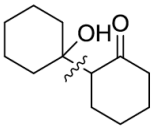

Figure 14 Dissociation of symmetric molecules

Some common synthons and their synthetic equivalents are shown in figure 15. Synthetic equivalents are the species that are used in the reaction instead of synthons due to better stability.

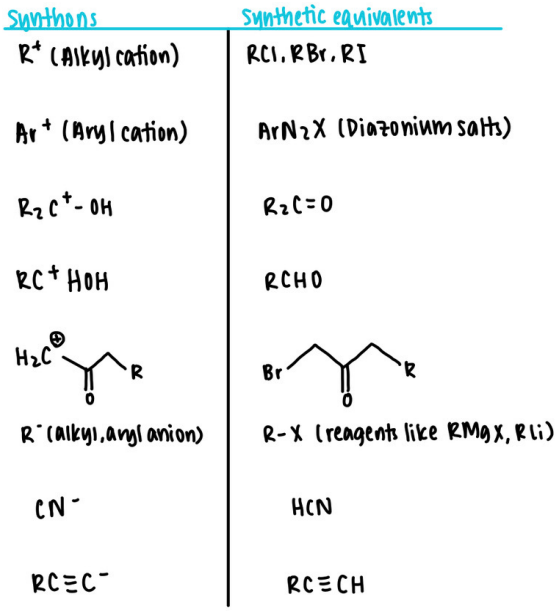

Figure 15 Common synthons and their synthetic equivalents

5.2 Example:

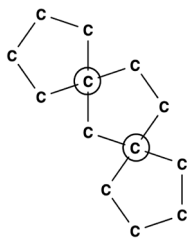

Figure 16 Three five-carbon-rings with two common ions

For a 13-carbon example shown in figure 15, it is easy to recognize the two common atoms C's, circled in the graph. There are three five-carbon rings as synthons by dissociating the target molecule into three parts through the common atoms.

6. Further examples of retrosynthesis

Retrosynthesis of 3 membered rings has mainly two approaches [5]:

1. One bond disconnection to make a long carbon chain.
2. Disconnection by removal of an atom. The atom removed can be either oxygen (epoxide) or an R group. As shown in Figure 17 below.

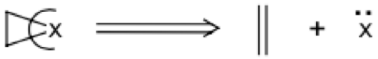

Figure 17 Disconnection by removing an atom [5]

For retrosynthesis involving the removal of oxygen, as shown in Figure 18, peracid RCO3H is used as the synthon for O. This is a stereospecific reaction.

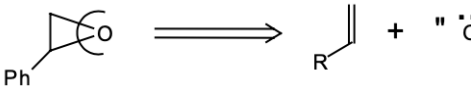

Figure 18 Disconnection by removing an oxygen atom [4]

An example of this approach is shown in Figure 19:

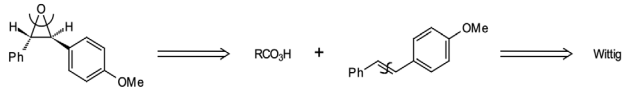

Figure 19 Example of retrosynthesis removing the oxygen atom [5]

For retrosynthesis involving removing an R group, Figure 20 is the common route.

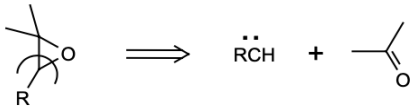

Figure 20 Example of retrosynthesis removing the R group [5]

Then this work will discuss two specific retrosynthetic routes in this case.

1) Using sulfur ylid as a synthon for CH2 as shown in Figure 21. Reacting with aldehydes or ketones can form the epoxide.

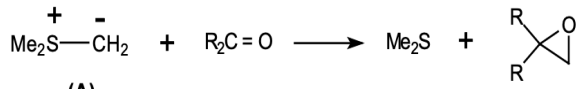

Figure 21 Pathway using sulfur ylid [5]

2) Darzen's condensation

Synthesis using Darzen's condensation is shown in Figure 22.

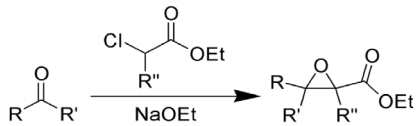

Figure 22 Darzen's condensation [5]

Retrosynthesis using this pathway is shown in Figure 23:

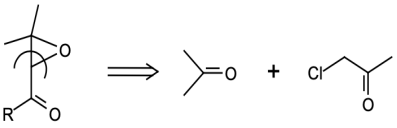

Figure 23 Retrosynthesis pathway using Darzen's condensation [5]

7. Another approach for retrosynthesis

Apart from the common atom approach, which is to locate the common atoms in two identical parts of a molecule, especially between rings, there are other ways to proceed with retrosynthesis. What if the molecule is not symmetrical, or how about the other nonsymmetrical parts of a macromolecule?

The work might focus on a step that is currently the easiest and probably disregard some duplicated steps while forming a molecule because they should be accomplished simultaneously. For example, when trying to synthesize 1-hydroxytasinine, the last stage of the reaction is shown in Figure 23.

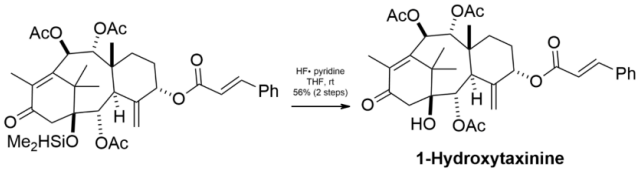

Figure 23 The last stage of the reaction [7]

This step substitutes (CH3)2HSi+ to H+, and there might have an intermediate along the process. Since there is only one hydroxyl group in the molecule, it is comparatively easier to make a sufficient amount rather than having other constitutional isomers as byproducts.

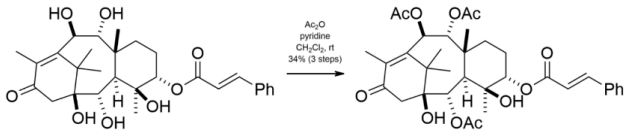

Figure 24 Several steps before the last reaction [7]

Figure 24 shows several steps before the last one, which is the procedure of substituting three H+ with Ac+ ions simultaneously, which is time-efficient.

8. Glossary

Synthon: A generalized fragment, usually an ion, produced by a disconnection.

Target Molecule (TM): The molecule whose synthesis is being planned.

Disconnection: breaking a carbon-carbon bond into smaller fragments.

Functional group interconversions (FGI): changing one functional group to another.

Synthetic equivalent: A reagent that carries out the function of a synthon that cannot be used by itself.

9. Conclusion

This work explains the basic ideas and concepts when doing retrosynthesis analysis. In addition, terminologies

and retrosynthetic analysis methods, such as synthons and FGI, are described and defined. Hopefully, this important intellectual tool can be further utilized in synthetic organic chemistry after reading the above. Especially in the medicinal industry, drug discoveries rely on synthesizing specific molecules that demonstrate particular abilities.

Reference

[1] E. J. Corey, X-M. Cheng (1995). The Logic of Chemical Synthesis. New York: Wiley. ISBN 978-0-471-11594-6.

[2] R. Singh, G. Gambhir, Pathshala. (n.d.). *Che p14 M24 etext - epgp.inflibnet.ac.in*. https://epgp.inflibnet.ac.in/epgpdata/uploads/epgp_content/S000005CH/P000669/M014013/ET/1456913012CHE_P14_M24_etext.pdf

[3] B.König, (2008), Strategy of synthesis; http://www-oc.chemie.uni-regensburg.de/OCP/ch/chb/oc5/Strategy_in_Synthesis-08.pdf

[4] Warren, S. (n.d.). Designing Organic Syntheses A Programmed Introduction to the Synthon Approach. JOHN WILEY & SONS.

[5] COREY, E. J. (1967). General methods for the construction of complex molecules. *The Chemistry of Natural Products*, 19–37. https://doi.org/10.101/b978-0-08-020741-4.50004-x

[6] Inoue (2019). 1-Hydroxytaxinine. chemistrybydesign.oia.arizona.edu/app/#1-Hydroxytaxinine%20(Inoue%202019)