

Introduction to Retrosynthesis: Approach to Do Disconnection and Practical Methods to Deal with Complex Molecules

Xinyu He^{1,*} and Yuchen Chang²

¹Sendelta International Academy, Shenzhen, 518108, China

²The College of Arts and Sciences, University of Washington, Seattle, 98195-4550, U.S.A.

Keywords: Disconnection, Synthons, FGI, FGR, FGA, Dioxygenation Patterns, Synthetic Equivalent.

Abstract: Retrosynthesis analysis is one of the most important ways to do full synthetic routes designing. All of its concept mainly include disconnection approach and synthons analysis respectively. In this passage, we want to introduce some of the effective approaches to do disconnection and explain the concept to the people who want to learn organic synthesis in the future or interested in this area. All about how to do disconnection, how to find and use synthetic equivalent, how to deal with one and two groups, Electrocyclic, and Illogical disconnection, how to use FGI, FGA, FGR to make those molecules which cannot be disconnected become possible, and how to deal with Dioxygenation Patterns etc. Also, at the end, we will discuss some examples that could enhance the memories.

1 INTRODUCTION

Imagine you're gluing bricks together when you see a photo of the finished artwork, and that's the end product. What is the first thing that springs to mind? The answer is to locate the materials you require and tie them according to the instruction book. Yes, this is the conventional technique, as well as how we normally think about synthesis. Retrosynthesis, on the other hand, will take the end product and imagine them into the fragment we already have, and it is similar to the 'finding materials' step we discussed before. Retrosynthesis may be thought of in this way at its most fundamental level.

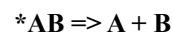
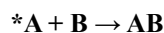
Retrosynthesis, often known as "the disconnection approach," is an analytical process in which a targeted organic molecule is deconstructed or fragmented to obtain starting material, or "Synthon". For a long term, many people have used this train of thought to design their way to synthesis, but there was not a clear definition. However, in 1964, Prof. Elias J. Corey, who was awarded Nobel Prize in chemistry due to his great contribution to synthetic organic chemistry. He was the first to formalize 'Retrosynthesis' this concept in his book 'The Logic of Chemical Synthesis' (Tutor, 2020; Corey, 1995). It provided different ideas to synthesize single and complicated target molecules.

For some extremely complex molecules, the basic goal is to generate precursors that correspond to available starting materials. In other words, retrosynthetic analysis is directed towards molecular simplification. Often, a synthesis will have more than one possible synthetic route. Retrosynthesis is well suited for discovering different synthetic routes and comparing them logically and straightforwardly.

Retrosynthetic analysis is a problem-solving technique for transforming the structure of a synthetic target molecule to a sequence of progressively simpler structures along a pathway which ultimately leads to a simple or commercially available starting material for chemical synthesis (Corey, 1995; Corey, 1988; Retrosynthetic analysis).

Take letters as examples:

Forward Synthesis Retrosynthesis



In order to get **AB**, **A** and **B** should be found first.

Furthermore, retrosynthetic analysis is particularly effective because there are so many different intellectual paths to pursue (Wang, 2022). By accessing its multiple possibilities of approaching routes, the most cost-effective, environmentally friendly, and concise path will be selected (Dmitrii A

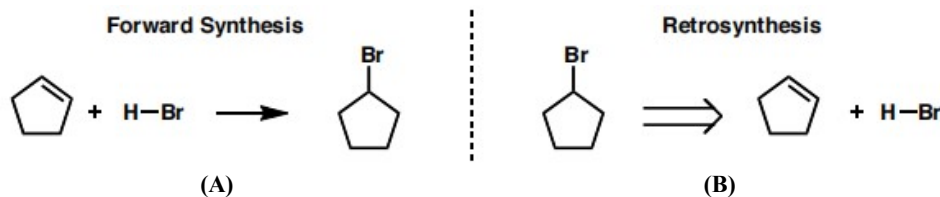
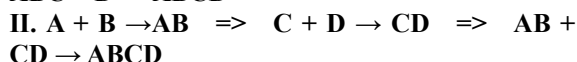
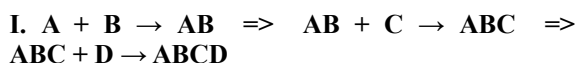


Figure 1: The example of the forward (A) and reverse (B) process of the hydrobromination of olefins.

, 2020). Take letters as examples again. In order to gain molecule **ABCD**, several ways to synthesis can be taken:



...

As can be seen, there are many ways to arrange and design paths to synthesis. Only thing need to be considered is comparing the advantages and disadvantages of each way, and select the best way to be taken.

2 STARTING TO DISCONNECTION

In the beginning, we mentioned that we may consider binding bricks as the way to synthesis. Now, we begin to disconnect some real molecules.

Disconnection means 'breaking' the bond of a molecule to generate simpler fragments. Actually, when disconnection occurs, instead of real bond breaking, it is more like brainstorm and a process in your head.

More simply, considering the analogy of a simple game, cutting things into pieces. The hexagon below shows that disconnection occurs, this shape is cut into some fragments as the right example denoted by the arrow. In figure 2, the two sticks are broken in the middle, and two lines are made up and down. There are several ways to cut this shape into pieces. But we need to obey the rules that keep the number of points at 6.

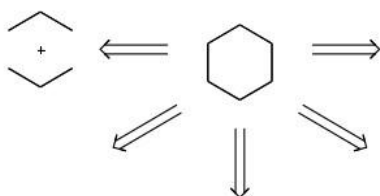


Figure 2: Showcase of a simple game.

In a similar way. Consider that this hexagon is a genuine Cyclohexane. Now you must recognize that more rules or restrictions have appeared. Following the same procedure to break the C-C bond. However, if you want to weaken the C-H bond, you must keep the skeleton formula's amount of invisible hydrogen constant. It's similar to breaking bonds, yet they're fundamentally different. As **Figure 3** shows, now consider a relatively complex molecule. In (1,1), we know a reaction to produce the molecule, then we can also do disconnection on it. But this is not the only way to produce the target molecule or the only way to disconnect. Thus, this lead us to find or try more reaction mechanisms in a long term. The more we know, the more simply the synthesis processes.

FXP 1:

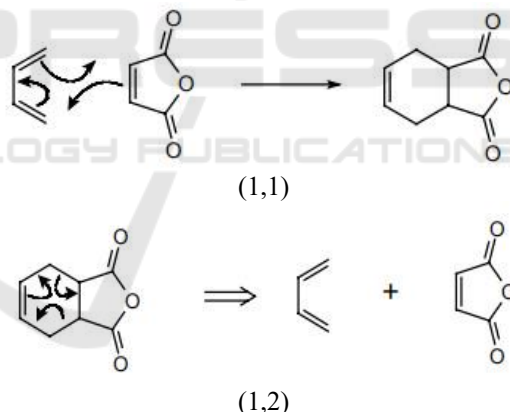


Figure 3: Reaction procedure.

3 APPROACH TO DISCONNECTION

3.1 One Group Disconnection

By now, we know how to do disconnection. Thus, we now need to consider how to allocate the electrons when bonds are broken. This is what we called bond cleavage. We have 3 situations to consider---Heterolytic cleavage (Left-side), Heterolytic cleavage(right-side), and homolytic

cleavage. In **Figure 4**, Heterolytic cleavage (right-side), the arrows show the direction of the electron's movement. When we break bonds, the two electrons are removed from the left-side molecule and pulled to the right-side molecule, making the left-side molecule cationic and the right-side molecule anionic.

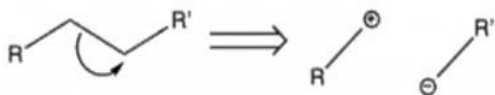


Figure 4: Heterolytic cleavage (right-side).

In **Figure 5**, there's another way to assign electrons. And it is similar to **Figure 4**. Two electrons are taken and pulled into the left-side molecule and make it anionic, and the other side becomes cationic.

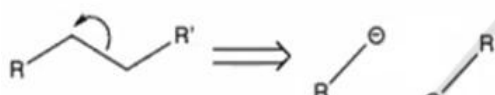


Figure 5: Heterolytic cleavage (left-side).

Otherwise, all of the above are synth ions, except for the final one. It is homolytic cleavage. When we break a bond, the electrons split and become free radicals on both sides. However, it is less common to make free radical (**Figure 5**) when bond disconnecting happens.



Figure 6: Homolytic cleavage.

When we finish our retrosynthesis analysis, it's not easy to find all of the materials we need, especially for some very complex molecules. So, how do we go on with our work? Normally, we will substitute other groups for the synthons we desire, and these groups are referred to as synthetic equivalents. A reagent that performs the function of a synthon but cannot be utilized itself, usually due to its instability.

For example:

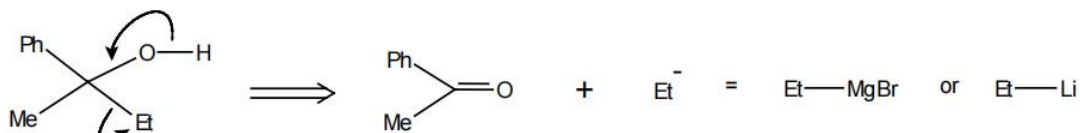


Figure 8: Example of retrosynthesis.

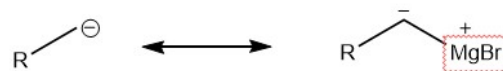


Figure 7: Equivalent function using Grignard reagent.

In **Figure 7**, it is obvious to see the carbon on the left is negative. We always use Grignard reagent to show an equivalent function when we need to do synthesis of some complex molecules. Since it shows a negative property in the middle.

In this example, we use the synthetic equivalent of the anion, the Grignard reagent or alkyl lithium, in this case because none of the stable anions are available. And in real retrosynthesis, when we talk about Et as a synthon, we're talking about EtMgBr or EtLi. Cause it will lead us to do disconnection and the whole synthesis process.

3.2 Two Group Disconnection

So far, we finished the one group disconnection, then how can we deal with the molecules with more functional groups? We must now consider the disconnection of the two groups, and I will demonstrate how to do so using the examples below. For example, in **Figure 9**, we have ketone and alcohol groups, and there are a lot of ways to do disconnection. The way we are easy to find is breaking the carbon in the middle and separating it into two parts. One is ketone and another is alcohol.

Nevertheless, we have a better way to approach it. The oxygen's delocalized electrons will offer us a more appropriate approach to detach. Then we obtained a positive and a negative synthon which is more common to find.

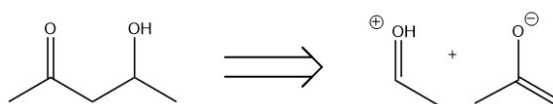


Figure 9: Two group disconnection.

Furthermore, we need to know what is **FGI**, **FGA** and **FGR**. And all of these are significant methods to do disconnection.

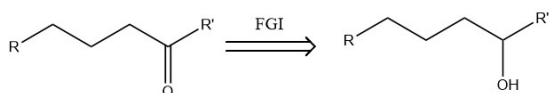


Figure 10: Example of FGI.

FGI fully spelled **Functional Group interconversion**. For this step, we always change our functional group to others in order to make sure we can disconnect the bond successfully. In **Figure 10**, the Carbonyl group was converted to an Oxhydryl group. And allowing the disconnection to occur.

FGA is the shorthand of **Functional Group Addition**. In real synthesis designing, it is sometimes necessary to add a functional group in order to enable the interchange of functional groups or later cut-off.

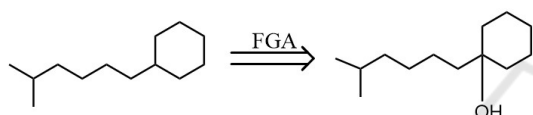


Figure 11: Example of FGA.

FGR is an abbreviation for **Functional Group Removal**. It means remove a functional group from a molecule. No like **FGI**, we need to remove some certain functional groups and make sure we can do disconnection. For instance, in **Figure 12**, we cut off Br group by doing **FGR**.

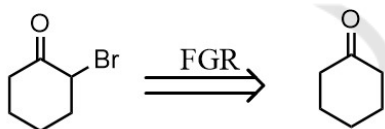


Figure 12: Example of FGR.

3.3 Electrocyclic Disconnection

In this part, I will explain to you a very important reaction mechanism that can dispose the complicated molecules effectively. It is **Diels - Alder Reaction**. And when we do retrosynthesis, it is the typical 4+2 disconnection.

Figure 13 below, it is the basic way to display how the process was. The left diene gets 2 double bonds, and the dienophile only has one double bond. As the arrow denotes, electrons from the dienophile attack the carbon (1) and make a bond between the diene and dienophile over there.

The double bond on the top of the diene (between C (1) and C (2)) will further relocate to carbon (2) and (3), then the double bond below will

connect with dienophile. All the six π bonds will shift at the same time, and eventually formed cyclohexene.

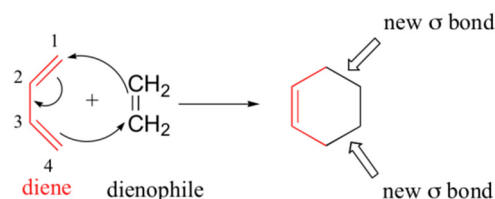


Figure 13: Basic process of electrocyclic disconnection.

3.4 Illogical Disconnection (Connection)

Illogical disconnection is not a common way to do disconnection, since it is making a bond rather than breaking a bond. The reaction below is a representative example to introduce this method. Using the double bonds on the aldehydes and combining them to form a single double bond and abandon the oxygen. Finally, we'll get cyclohexene. We may achieve this by using ozone, and the reaction is known as Ozonolysis.



Figure 14: Example of illogical disconnection.

3.5 Dioxygenation Patterns

In this part, it is clear to see that the main purpose of this process is dealing with 'oxygen'. We will face to a following of reactions, and I think the basic synthetic way is to simplify problems into the model we have worked out in the past.

3.6 1dioxygenation Patterns

As shown, once we see these patterns, we should convert them into OH-groups by using FGI, then OH-groups could disconnect to olefin.

3.7 2dioxygenation Patterns

For these molecules, FGI is used and we will have a carbonyl and a hydroxyl group. Then we can repeat the steps (Two groups disconnection) we mentioned before.

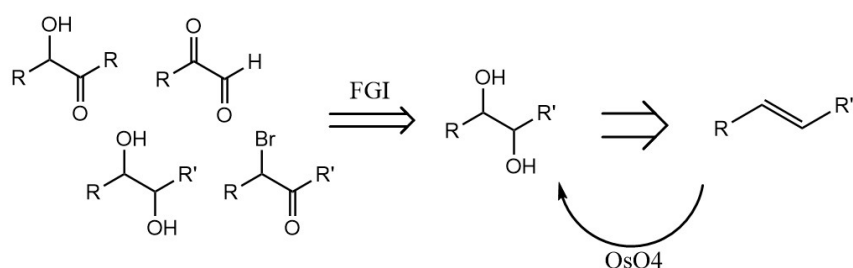


Figure 15: 1dioxygentation patterns.

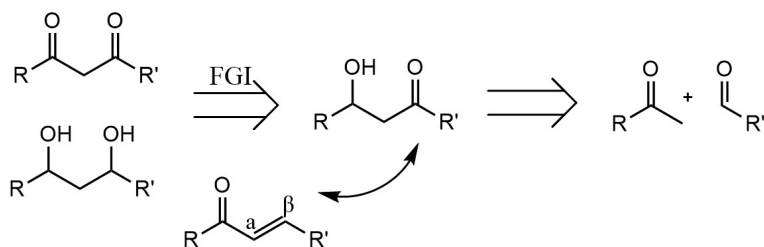


Figure 16: 2dioxygentation patterns.

3.8 3dioxygentation Patterns

In this part, we need to convert the groups into two carbonyl groups, and disconnect them into two parts.

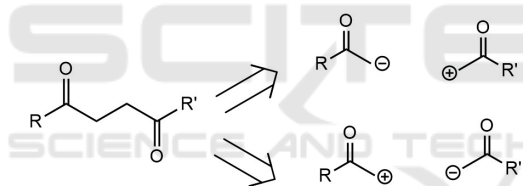


Figure 17: 3dioxygentation patterns.

3.9 4dioxygentation Patterns

For these molecules, we want to make di-carbonyl groups, thus we need to use two carbonyl groups. And we have equivalent molecules. Using the conjugate addition mechanism, when these two molecules come into contact, the anion will attack the double bond and form the di-carbonyl compound.

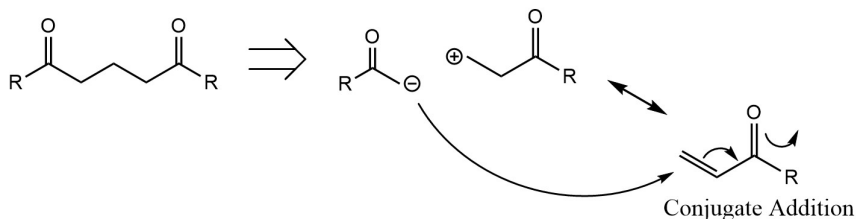


Figure 18: 4dioxygentation patterns.

3.10 5dioxygentation Patterns

For these molecules, we need to convert into alkene with two carbonyl groups and repeat the step (Illogical Disconnection) again.

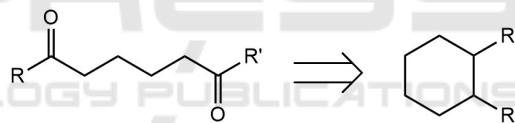


Figure 19: 5dioxygentation patterns.

4 FOCUS ON MAXIMIZING SIMPLIFICATION (FOCUS ON SYMMETRY)

Efficient and concise synthetic routes based on the symmetry of the molecule are gaining widespread attention, and such synthetic routes can be two-directional synthesis.

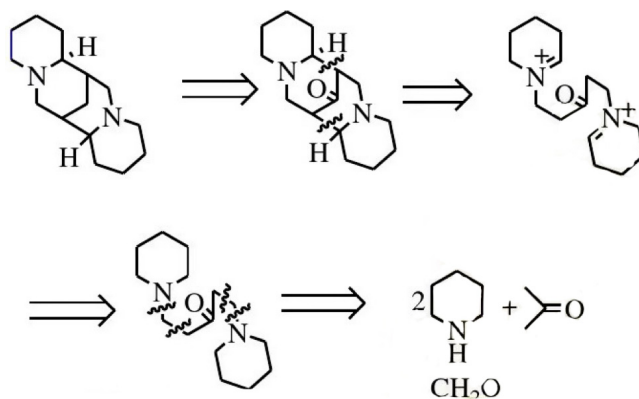


Figure 20: Synthetic routes of two-directional synthesis.

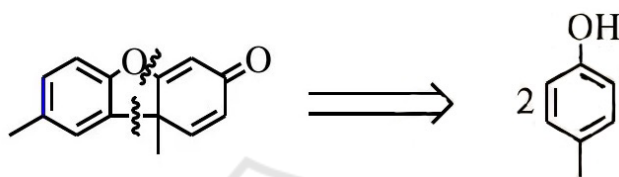


Figure 21: Process of reaction.

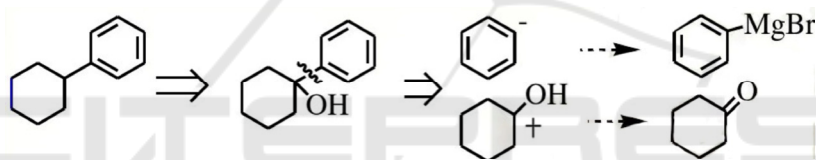


Figure 22: Process of reaction.

eg: sparteine has symmetry by introducing the carbonyl group on the central methylene and then symmetrically using the inverse Mannich cut on both sides to highly simplify the molecule. This yields three basic raw materials: piperidine, formaldehyde, and acetone, all of which are synthesized by the classical standard reaction.

Some target molecules do not have symmetry per se, but have potential symmetry, and after certain inverse synthetic transformations, a symmetric molecule or a symmetric synthetic route can be obtained, thus simplifying the synthetic design. For example, there is no symmetry factor in Pummerer's ketone molecule, but the two radicals obtained after the cut is from the same precursor.

Addition of auxiliary functional groups followed by cleavage (FGR).

Some target molecules can be cut off only after the addition of appropriate functional groups to find the correct route of synthesis. In the example figure 22, there is no functional group, and when a hydroxyl group is introduced into the cyclohexyl group, it can be cut further.

In the retrosynthetic transformation of the target

molecule, it is required that some necessary structural unit exists in the target molecule, and only when such a structural unit exists or such a substructure can be generated, can the target molecule be effectively simplified and easy-to-access starting materials be deduced, such as the following structural units of A, B is the basic retron of Diels-Alder reaction, Robinson formation, respectively.

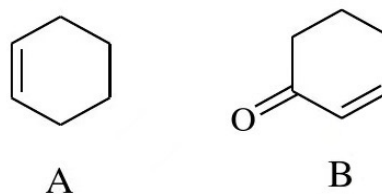


Figure 23: Structural units.

The core problem of retrosynthesis analysis is transformation, and inverse retron and synthon are two aspects of this core problem, the former is the necessary structural unit for transformation, and the latter is the structural unit to be obtained by transformation.

5 CONCLUSION

The trend of organic synthesis is not to blindly pursue new compounds, but to design and synthesize compounds that are expected to have excellent properties or have great significance. The retrosynthetic analysis method takes a complex synthetic problem and decomposes it into several simple synthetic problems by dissecting it step by step from tedious to simple through the inverse method, and then forms a synthetic route from simple to complex molecules. It makes many difficult applications possible and saves research costs, such as the application of drug synthesis. Such a mode of thinking can also be applied to different disciplines, such as synthetic aperture radar for meteorological and atmospheric motion studies, mathematical studies of inverse matrices, economic reverse logistics.

REFERENCES

- Dmitrii A. Shabalin, Elena V. Ivanova, Igor' A. Ushakov, Elena Yu. Schmidt, Boris A. Trofimov, Retrosynthetic Analysis of α -Alkenyl- β -Diketones: Regio- and Stereoselective Two-Step Synthesis of Highly Arylated Representatives from Acetylenes, Ketones, and Acyl Chlorides, *The Journal of Organic Chemistry*, Volume 85, Issue 13, 2020, Pages 8429-8436.
- E. J. Corey, X-M. Cheng (1995). *The Logic of Chemical Synthesis*. New York: Wiley. ISBN 0-471-11594-0.
- E. J. Corey (1988). "Retrosynthetic Thinking - Essentials and Examples". *Chem. Soc. Rev.* 17: 111-133. doi:10.1039/CS9881700111.
- Retrosynthetic analysis https://www.chemurope.com/en/encyclopedia/Retrosynthetic_analysis.html#_note-0/
- Tutor, 13 Jul, 2020, Retrosynthesis, <https://onlineorganicchemistrytutor.com/retrosynthesis/>
- Wenlong Wang, Qilei Liu, Lei Zhang, Yachao Dong, Jian Du, RetroSynX: A retrosynthetic analysis framework using hybrid reaction templates and group contribution-based thermodynamic models, *Chemical Engineering Science*, Volume 248, Part B, 2022, 117208