

Introduction to Retrosynthetic Analysis

Shangyou Chen^{1,*}, Zhangsheng Liu²

¹Tianjin Foreign Languages School Affiliated to Tianjin Foreign Studies University, Tianjin, 300230, China, tfls_jerry2020@126.com

²Tongwen School (Jiaxing), Jiaxing, 314050, China, 15821528205@139.com

tfls_jerry2020@126.com

Abstract: The purpose of this paper is to act as a fundamental guideline covering the starting retrosynthesis concepts and to incorporate some of the essential skills and knowledge for new starters in designing a retrosynthesis plan. The retrosynthesis approaches such as fine tuning and various group disconnections are presented before showing a full and detailed retrosynthesis analysis example.

Keywords: Fundamental guideline, Concepts, New starters.

1. Introduction

Retrosynthetic Analysis is a fundamental analysis used in organic synthetic chemistry. It was first suggested by Professor E J. Corey. Because of his incredible discovery, Professor Corey won the Nobel Prize in chemistry in 1990. After that, organic chemists use the method called retrosynthetic method to approach their target molecules. People used to extract drugs and pigments from plants and marine animals, but in a very small amount. For example, a pigment called Tyrian purple was extracted from the yellow slime of some of the sea snails in the family Muricidae. In ancient times, extracting this dye involved tens of thousands of snails and substantial labor, so as a result, the dye was highly valued. However, the main chemical constituent of the Tyrian dye was discovered by Paul Friedländer in 1909 to be 6,6'-dibromoindigo[1] as shown in Figure 1. Chemists can do a retrosynthesis on this molecule and synthesize commercially. An efficient protocol for laboratory synthesis of 6,6'-dibromoindigo was finally developed in 2010 and the prize finally fell down. [1,2] The use of retrosynthesis in organic synthesis evidently improved people's quality of life, by finding more paths to approach a molecule thus decrease the expense of buying it.

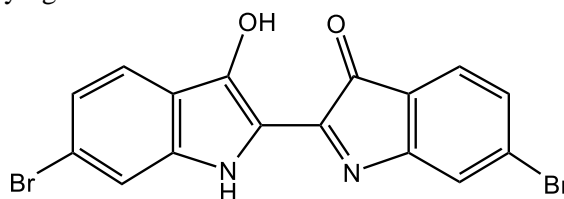


Figure 1. 6,6'-dibromoindigo.

The hardest part in designing a retrosynthesis pathway in the first place is making a decision to choose a site for breaking a bond, or so-called the bond-disconnection for our Target Molecule in order to select a reagent or approach to originate the "products", but instead every newly derived step before ending should be regarded as the last step in our retrosynthetic strategy for our target molecule.

Before introducing the various approaches in retrosynthesis, some basic knowledge should be presented first in order to pave the ways for reasoning some retrosynthesis functions more readily.

2. Basic Information

2.1. Bond Polarity

It may appear to be familiar to folks, and indeed, it's depicting a relative extent to which a bond connecting two species has a polarity that is apt toward either side of the atom.

In addition, the polarity is commonly given rise by the electronegativity of the directly bonded atoms, which is a relative strength for the nucleus of an atom to attract electrons in covalent bond [3]. The polarity, however, can result in a partial negative sign on the more electronegative atom and the partial positive sign on the less electronegative one. The concept of polarity is to facilitate the recognition of the idea that as the bond is polarized, the more electronegative atom will strive to obtain the two electrons from covalent bond. In most of the cases, nitrogen, oxygen, and the halogens have the partial negative sign assigned.

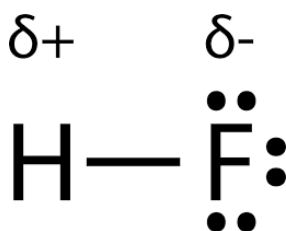


Figure 2. Partial positive and negative signs labeled in a hydrogen fluoride molecule.

2.2. Electron donating groups and electron withdrawing groups/ EDG and EWG

Electron donating groups are usually the substituents that give away or really donate their electron density to the nearby carbon atom through resonance into a conjugated pi system. Typically, in a carbon cation structure as the substituents, the reason for the donation of some of the methyl groups' electron clouds is due to the hyperconjugation, which is a phenomenon focusing on the electrons in between the C-H or C-C sigma bonds as the electron-rich methyl groups interact with the empty p orbital on a nearby carbon atom as the sigma bond rotates freely and then overlaps with the orbital to share some electron density [4]. The overall effect is stabilizing the structure and becoming more nucleophilic.

While EDG pushes electrons into the system, the electron withdrawing groups, the EWG, pulls portion of the electron clouds away from the system onto themselves, making the structure less stable without those electrons at their positions. The cause of electron-taking ability is due to the inductive effect, which develops partial negative on the more electronegative atom and a partial positive sign on the less electronegative atom to donate the fact that the electron density is being pulled through bond. The most representative substituents acting as EWG are hydroxyl group -OH, carboxyl group -COOH, and the strongest one trifluoromethyl group -CF₃ since it has three the most electronegative atoms attached to a carbon.

2.3. Leaving groups

A leaving group is an atom or functional group that can drop off the molecule that it connects with and remain with a lone pair of electrons [5]. In short, it performs the exact opposite to a nucleophile; instead of donating electrons to form a covalent bond, it breaks the bond with the molecule and takes the

electrons as it leaves. Due to the loss of electrons, the primary molecule is left missing electrons, becoming a cation. Moreover, the halogens earn their position in the leaving group as the 17 group elements are still stable in the form of halides with an additional charge assigned. The example of bromine in 2-bromopentane suffices to illustrate this concept easily.



Figure 3. Bromine as a leaving group detaches from the 2-bromopentane forming the carbocation intermediate.

However, fluoride ion couldn't become a member of the leaving group due to its small radius and also the concomitant high charge density. Owing to a large repulsion between electron inside a fluoride ion, such unpolarizable effect deters it from leaving the molecule promptly with electrons on the same side even though it's the most electronegative atom.

2.4. Resonance

Not only does the nature of a weak base decide what a leaving groups would be, but resonance also plays a role in the stabilization of a molecule by relocating the electrons or the movement of the electron density throughout the whole structure within a conjugated pi system. As long as the leaving group could be stabilized through resonance, it's generally a positive indication of its ability.

2.5. Oxidation level

Generally, three types of oxidation levels are usually incorporated, which are the alkane, alcohol, aldehyde/ketone, and the acid oxidation level. In terms of the difference among the three, it has to do with the number of electronegative atoms attached on the carbon atom. Alkane has structure in which only hydrogen atoms are attached; alcohol level has one electronegative; aldehyde/ketone has two; and finally, acid has three. What's worth noticing is that a double bond counts as two electronegative atoms. The significance of such categorization of these oxidation levels implicates the operations in retrosynthesis, and that is the transformations made among the same oxidation level is common and readily, for example in the case of a carboxylic acid and a chloroacetone: since double equals to two electronegative atoms, both molecule each has in total three electronegative atoms, or, the acid oxidation level.

2.6. Protective group

A protective group, or protecting group, is a functional group that is introduced during some reactions that the functional group chemists want to preserve would react with the reagents, and to avoid side products which is unwanted. Introducing a protective group needs two extra steps among the whole synthetic process: introducing process and removal process. These functional groups are common that they should have a known yield of the introducing process and the removal process. Here are some other strategies that a idealized protective group should have: the reagent for introducing should be cheap and easily got; the intermediate formed should be mild and stable, and won't cause any side reaction; won't introduce any chiral center to the molecule; and finally, the product should be easily removed/isolated.

Here is a classic example of introducing a protective group, called acetal protection, as shown in Figure 4:

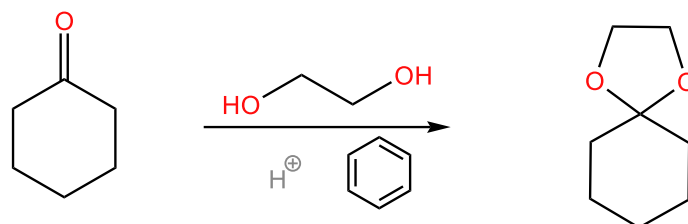


Figure 4. Acetal Protection on Cyclohexanone.

The reaction starts with a ketone and ends with a group of compounds called acetal with acid catalyzed and benzene as the solvent. This process is a reversible, and the equilibrium lies highly towards the left side. But how can this unstable reaction be a classic example? Actually, the side product of this reaction is H_2O , which is a easily removable product, and then according to the Le Chatlier's Principle, the removal of the product will shift the equilibrium to the product part. This is the reason why we are using benzene as a solvent: it's insoluble in water. Reversibly, by adding extra H_2O to the solution, the reaction will shift to the left which we get the protected group back.

The carbonyl group is protected during some particular reactions. Like in the reaction where you want a ester being reduced to a 1° alcohol but there is a ketone in the molecule. According to the table of pKa Values for a Few Common Organic Acids, ketones have the pKa value between 20-24, and that of esters are between 24-25, which tells us that ketones are always more reactive than esters, so if an ester wants to react, a ketone goes before it. As an example, in the reaction with LiAlH_4 , both ketone and ester can be reduced to an alcohol. And here is where we should use the acetal protection.

3. Retrosynthesis approaches

Generally, a typical retrosynthesis analysis is done by means of **bond-disconnection** and **Fine Tuning** [2].

For a bond disconnection, it's a conceived as a cleavage of a bond resulting in the situation of synthons, which are the idealized fragments from a disconnection that could help with the synthesis plan. The imagined bond cleavage corresponds to the reverse of the real reaction in synthesis plan [6].

3.1. Synthon and Reagent

A reagent performs as a species that possess the equivalent functionality to that of a synthon, since a synthon is idealized and is usually less promptly or readily to be obtained, in fact, it should be substituted with something that is similar, which is a reagent and also called as a **synthetic equivalent**. The synthetic equivalent carries out the features of a synthon since the synthon couldn't be created often. In addition, synthons are with a positive/negative charge on the atom that experiences the bond disconnection. Furthermore, a reagent is expressed as equating to a synthon by a triple line \equiv . In the case of a 2,4-Dichlorophenoxyacetic acid shown in figure 5, the synthons, the intermediate molecules with a charge on the oxygen and carbon atom, can be replaced respectively by 2,4-dichlorophenol and ethanoic acid.

2,4-Dichlorophenoxyacetic acid as the target molecule and the imagined synthons

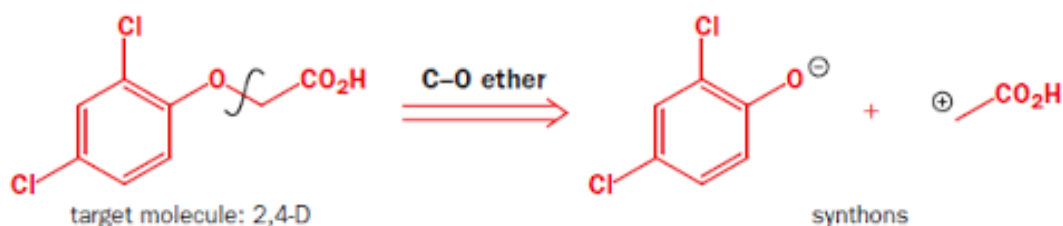


Figure 5. 2,4-Dichlorophenoxyacetic acid.

3.2. Fine Tuning

For Fine Tuning, we have several methods to apply for. The first one is called the Functional Group Interconversion, the FGI, it is a process of transforming a functional group into another by either addition, substitution, elimination, oxidation, or reduction. This operation is frequently exploited and applied to numerous structures of molecules when only writing a functional group for another so that the disconnection becomes possible or more readily. For example, cyclohexanone can be converted to cyclohexanol via the FGI process, as shown below in Figure 6.

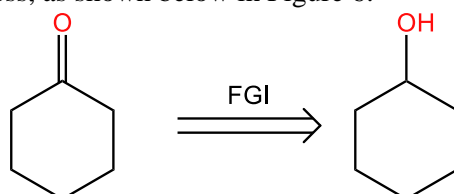


Figure 6. FGI Turning Ketone to Alcohol.

Besides FGI, we can do another process to simplify a complex molecule: the Functional Group Removal, or the FGR. As it is mentioned in the word, this is a process of removing functional groups in a molecule. A complex molecule always have various functional groups so FGR can be done differently when treating one molecule. As an example, 2-bromocyclohexan-1-one, have a halo group and a carbonyl group in it, without the consideration of the stereoselectivity of this chiral molecule, it can be converted in two ways as shown in Figure 7. The first one is using cyclohexanone for a halogenation reaction, the other one is using bromo-cyclohexane and do some reactions to introduce a ketone to the molecule to finally make 2-bromocyclohexan-1-one, our target molecule.

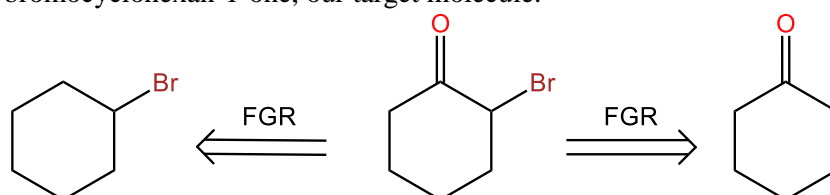


Figure 7. Two FGR Turning Pathways on 2-bromocyclohexan-1-one.

There is a last strategy of fine tuning, called the Functional Group Addition, or the FGA. Comparing to the two strategies above, it might be a little confusing that the purpose of the whole process of retrosynthesis is simplifying and the FGA is adding functional groups. However, sometimes a complex molecule would be more abundant, or easier to get in real life. For example, cyclohexa-1,3-diene, is our target molecule, and is not that abundant in the real life; but when we do a FGA as shown in Figure 8, we get benzene, which is a really common product whether using cyclic polymerization with ethyne or using steam cracking with petroleum.

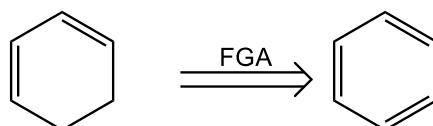


Figure 8. FGA Turning Cyclohexa-1,3-diene to Benzene.

3.3. Disconnections

3.3.1. One-group Disconnection If the target molecule is monofunctional (having 1 functional group only), the disconnection process is classified as a one-group disconnection.

For example, pentan-3-ol in Figure 9,

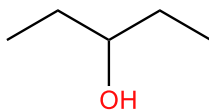


Figure 9. Penta-3-ol.

If we directly breaking the C-C bond on either side of this symmetric molecule, we get this in Figure 10,

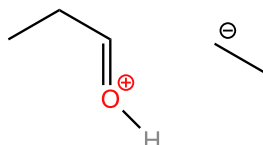


Figure 10. Direct Breaking Resulted in Synthons.

They are synthons, and we need the reagent. Their synthetic equivalents could be these in Figure 11,

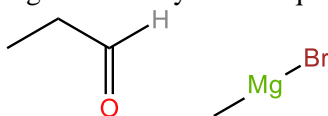


Figure 11. Synthetic Equivalents of Synthons in Figure 10.

3.3.2. Two-group Disconnection

If the target molecule is difunctional (having 2 functional groups), the disconnection process is classified as a two-group disconnection.

There are many ways of disconnecting bonds, for instance, in Figure 12

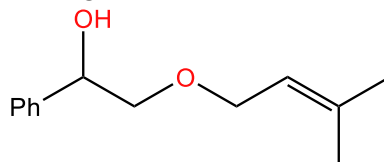


Figure 12. C₁₃H₁₈O₂ [7].

If we disconnect the C-O bond on the side close to the alkene group, it will break into this as shown in Figure 13,

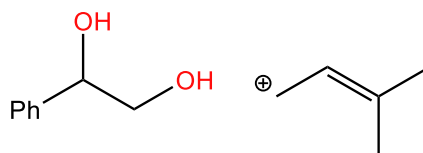


Figure 13. First Pathway of Direct Breaking Resulted in Synthons[7].

However, apparently the molecule on the left has 2 alcohol groups which would cause chemo-selectivity problem that each OH group would react.

But what about disconnecting the C-O bond on the side close to phenyl group? If we do this second disconnection, it will break into this as shown below in Figure 14,

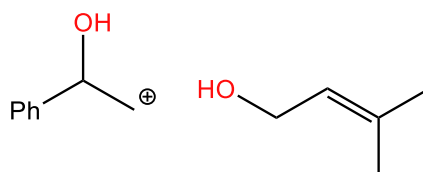


Figure 14. Second Pathway of Direct Breaking Resulted in Synthons[7].

This one is more workable that we avoid a potential chemo-selectivity problem by doing a two-group disconnection.

4.Examples of Retrosynthesis

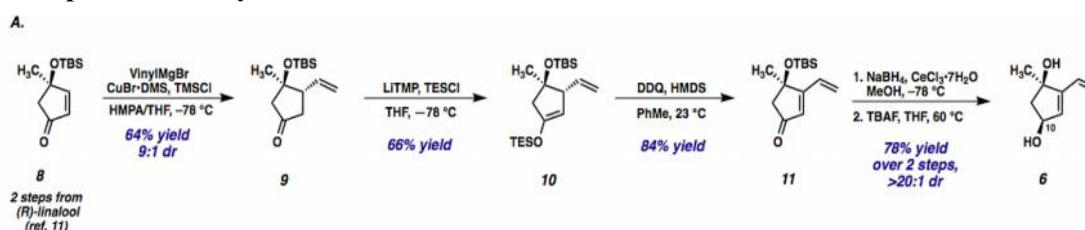


Figure 15. Forward Synthetic Analysis [8].

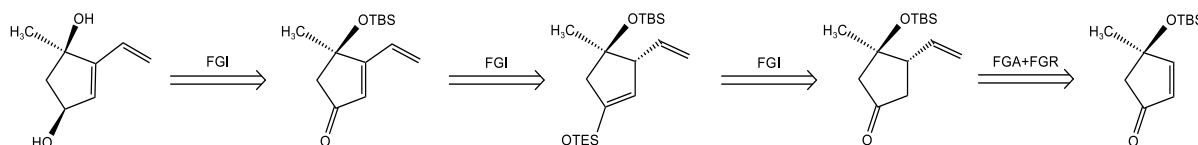


Figure 16. Retrosynthetic Analysis.

This is a synthetic process for a reagent from the total synthesis of a compound called Scabrolide A, and we will use this as an example explaining what each step contributes to the total synthetic process.

The first step in the retrosynthesis (in Figure 16), or the last step in the forward synthetic analysis (in Figure 15), the hydroxyl group in the lower portion of the molecule is converted to a ketone via a FGI process. At the same time, the upper hydroxyl group is converted to a OTBS group also through a FGI process.

On the second step, the aldehyde group on molecule 11 is converted to a single bond attached to a TESO group and an adjacent alkyl group, which is also done by FGI.

When doing retrosynthetic analysis on molecule 10, we can see that the TESO group is replaced with an aldehyde group and the electrons previously in the double bond inside of the ring is transferred to the CO double bond. But wait, we did a FGI to convert the ketone to TESO in the previous analysis, but in this step, we reassemble the ketone. This is because the reagents used in the synthetic process of molecule 10 can turn the ketone into some functional groups that would make the following synthesis difficult to approach. So, the TESO group was just a protective group preventing complicated chemo-selectivity problem that may occur.

On the final step of retrosynthesis, we use a FGR and an FGA at the same time, which removes an alkene that attaches to the ring and adds an extra bond to the ring.

Up here, we are still left with this relatively complicated molecule 8, so we might do another retrosynthetic analysis on molecule 8, until it gets to a compound that has up to six carbon atoms and with one functional group, which would usually be an available starting material to get from suppliers.

[7]

5. Conclusion

The core in retrosynthesis is to deconstruct a chosen target molecule to be extended from sequences of plausible starting materials and reagents, and its ultimate goal is to source the most available and efficient retrosynthesis pathway. It's the exact opposite of a synthesis. As the first paragraph suggested, the concept of retrosynthesis is to produce product in a reverse way, and it solves the problem of high expense in purchasing a product [8]. Last but not least, it should be emphasized that creativity enable thought in retrosynthesis to flourish since one can figure out an optimum pathway in retrosynthesis after reading this article.

6. Acknowledgement

Shangyou Chen and Zhangsheng Liu contributed equally to this work and should be considered co-first authors.

References

- [1] Wikipedia contributors. (2022) Tyrian Purple., en.wikipedia.org/wiki/Tyrian_purple.
- [2] Beubenz, 2018, Organic Synthesis by Retro Synthesis. https://www.slideshare.net/ShariqueKhan22?utm_campaign=profiletracking&utm_medium=sssite&utm_source=ssslideview.
- [3] Chemistry LibreTexts. 2022. 3.4: Bond Polarity. https://chem.libretexts.org/Courses/Mount_Royal_University/Chem_1201/Unit_3%3A_Chemical_Bonding_I_-_Lewis_Theory/3.4%3A_Bond_Polarity
- [4] Hunt, 2022, Hyperconjugation. <https://www.chem.ucalgary.ca/courses/353/Carey5th/Ch04/ch4-3-2-1.html>
- [5] Valqui, M., | ChemTalk. 2022. What is a Leaving Group? <https://chemistrytalk.org/leaving-group/>
- [6] Clayden, J., Greeves, N., Warren, S. (2012) Retrosynthetic Analysis. In: (Eds.) Organic Chemistry, 2nd ed., Oxford University Press Inc., New York, pp. 703–711.
- [7] Hafeman, N., Loskot, S., Reimann, C., Pritchett, B., Virgil, S., Stoltz, B. (2020) The Total Synthesis of (–)-Scabrolide A. *J. Am. Chem. Soc.* 2020, 142, 8585–8590.
- [8] Liu, C., Blog.cambridgecoaching.com, 2022. The basics of retrosynthesis. <https://blog.cambridgecoaching.com/the-basics-of-retrosynthesis>