An Attempt Method for Developing New Synthetic Routes by Fusing Computational Chemistry and Chemoinformatics: Syntheses of Ethyl and Benzyl Methacrylates

Kenzi Hori,⁎ Hirotaka Sadatomi, Katsuhiko Okano, Michinori Sumimoto, Atsuo Miyamoto, Saori Hayashi, and Hidetoshi Yamamoto

⁎Department of Applied Chemistry, Graduate School of Science and Engineering, Yamaguchi University, Tokiwadai, Ube 755-8611, Japan, †Tokyo R&D Center, Daiichi Pharmaceutical Co. Ltd., Kitakasai, Edogawa, Tokyo 134-8630, Japan, and NARD Institute, Ltd., Nishinagasu, Amagasaki 660-8050, Japan.

(Received December 1, 2006; Accepted January 5, 2007)

In synthesizing target compounds, synthetic chemists search for and examine papers concerning syntheses to create their own synthetic routes. Synthetic route design system programs such as AIPHOS and EROS can also create many synthetic routes for targets compounds. However, it takes a long time to experimentally test all the routes created by these programs. Therefore, computational chemistry has been used to investigate mechanisms of chemical reactions which have already been examined experimentally. It is feasible to consider theoretical methods as effective in examining whether or not synthetic routes from chemoinformatics can be used practically for synthesizing target compounds. There are many difficulties associated with the aforementioned application, because theoretical calculations have to be done without information from experimental works. In the present study, we proved that this concept works well for synthetic routes of benzyl methacrylate, created from the KOSP and TOSP programs.

Key Words: computational chemistry, chemoinformatics, synthetic route, methacrylate

⁎ kenji@yamaguchi-u.ac.jp
Introduction

Synthetic route design systems (SRDSs) such as LHASA,1 EROS2 and AIPHOS3 were developed to create new synthetic routes for many compounds. The KOSP (Knowledgebase-Oriented Synthesis Planning) program,4–5 one of the AIPHOS family programs, is now commercially available and has been used practically to create novel synthetic routes. It is, therefore, possible to obtain synthetic routes for target compounds without having a detailed knowledge of organic syntheses.

Theoretical calculations are usually very powerful to investigate chemical reactions and have been applied to analyze mechanisms of organic and inorganic reactions.6–8 The transition state (TS) search is the key to understanding reaction mechanisms in detail, for example, why good selectivity for a target was achieved with a given compound. A similar approach should be useful for judging whether or not synthetic routes from SRDSs produce target compounds even though the routes have not yet been examined experimentally. As all the routes, which SRDSs produce, do not always lead to the formation of target compounds, this analysis can be done before beginning to synthesize target compounds. However, there are difficulties associated with doing such applications since theoretical calculations have to be done without information from experimental works. Moreover, if one can extrapolate yields from reactions for SRDSs targets, one can determine which synthetic route should be examined first.9

In the present study, we developed new synthetic routes for benzyl and methyl methacrylate 1a and 1c by combining computational chemistry and chemoinformatics as shown in Scheme 1. We adopted the relatively simple synthetic targets because this is the first real attempt that includes theoretical and experimental methods.10 Although synthetic routes for 1 are not very important for synthetic chemists, this is a very important first step to prove that the combination of computational chemistry, chemoinformatics and experiments is useful for developing new synthetic routes for target compounds.

In order to create synthetic routes for 1a and 1c, we used KOSP11 and TOSP (Transform-Oriented Synthesis Planning)12 programs. The former program produced seventeen routes and the latter three. Four routes in Scheme 1 were chosen from the routes created by KOSP and TOSP programs since it is rather easy to synthesize the reactants. Moreover, the routes are simple enough to perform experiments, which can be compared to results from theoretical calculations. The first, Route A, applies a Favorskii rearrangement of 3,3-dibromo-2-butanone 2 with alkoxide ions. Route B uses a Horner-Emmons reaction of 2-(diethoxyphosphoryl)propionate 3 and formaldehyde. Route C is a reaction between propionate and Eschenmoser’s reagents and the last is a simple SN2 type reaction that methacrylate ion attacks alkyl bromide at the back side of Br (Route D). The detailed mechanism of these reactions will be further discussed in a subsequent section.

Calculations and Experimental Methods

Density Functional Theory (DFT) Calculations

Geometry optimizations including TSs were carried out using the Density Functional Theory (DFT) calculations in the Gaussian03 program13 at the B3LYP/6-31G* or B3LYP/LANL2DZ level of theory.14–16 WinMOPAC Ver.217 and Jmol18 programs were used for molecular modeling. The intrinsic reaction coordinates ( IRCs) at the PM3 level of theory were also calculated to analyze mechanisms for Routes B and C. Vibration frequencies were calculated to confirm geometries for stable or TS structures.

Experimental Section

All chemicals were purchased as reagent grade and used without any further purification. All reactions were carried out under a nitrogen atmosphere. Tetrahydrofuran (THF) was dried over sodium metal and benzenophenone before use. The progress of reactions was monitored by TLC (thin layer chromatography) (silica gel 60 F254, Merck). Chromatographic purifications were performed with Wakogel C–200 (100–200 mesh). NMR spectra were recorded at 270 MHz for protons and at 67.8 MHz for carbon nuclei in
(CDCl₃) with a JEOL EX-270.

**Results and Discussion**

**Route A**

The first synthetic route includes a Favorskii rearrangement of 2. The detailed mechanism is shown in Eq. 1. An alkoxide ion extracts one methyl group hydrogen adjacent to the carbonyl group of 2, followed by the formation of 2-bromo-2-methylcyclopropanone 7 with a three-membered ring via a transition state 7(TS). According to the DFT calculations, the formation of the small ring is the rate-determining step with an activation barrier of 19.6 kcal mol⁻¹ as shown in Figure 1.

The addition of another alkoxide ion to 7 produces a tetrahedral intermediate 8, followed by the breaking of the small ring, which allows the Br⁻ ion to form product 1. The 8a(TS) is more unstable than 7a by only 8.2 kcal mol⁻¹ as shown in Figure 1. The inclusion of dimethylsulfoxide (DMSO) as a solvent is essential for geometry optimizations of 8a(TS) and 8a. The alkoxide ion prefers extracting a hydrogen in the ring to forming a tetrahedral intermediate 8a. The activation barrier for the extraction of a proton from 7 was calculated to be only 0.9 kcal mol⁻¹, which is much smaller than the corresponding value for 8a. As 9 has a highly distorted three-membered ring, it easily decomposes to something else. Therefore, we can conclude that Route A is not worth doing further experimental efforts.

**Route B**

The Horner-Emmons reaction uses phosphonates. Ando has already used an ethyl ester and an aldehyde in THF by a NaH extraction an activated hydrogen (Eq. 2). The benzyl ester was also used to prepare an olefin of Eq. 2 in the same solvent. However, we could not find any examples using simple formaldehyde to obtain the ethyl or benzyl acrylate targets in the present study.

According to the reaction mechanism of the Horner-Emmons olefination, the reaction of 3 and formaldehyde is expressed in Eq. 3. Formation of the anion 10 is expected to occur easily, and it reacts with formaldehyde to form 11. However, experimental results indicated that this process is not as simple as expected. Formation of 10 is highly dependent on bases as will be discussed in a subsequent section.

The intermediate 11a was calculated to be less stable than 10a by 3.2 kcal mol⁻¹ and the activation barrier for its formation from 10a was calculated to be only 3.7 kcal mol⁻¹ as shown in Figure 2. A minimum energy path calculation on the O-C bond in the middle of the geometry of 12a showed that the bond breaks very easily to form the final product almost without an appreciable barrier. These results indicate that 1 will be synthesized by applying the Horner-Emmons reaction of 3 and formaldehyde. Therefore, it is concluded that this is a feasible synthesis route for producing the target.

As the previous two experiments used NaH as the
base, we first examined this reagent and obtained only trace products. On the other hand, the usage of t-butoxide ion lead to the production of the target 1a with a yield of 47%, and this value was obtained without optimizing reaction conditions. Although the choice of the base is very sensitive to generating the anion intermediate, the theoretical conclusion is consistent with the experimental results.

**Route C**

According to the mechanism of Eschenmoser’s salts in the textbook, the synthesis route for 1 consists of three reactions as shown in Eq. 5. The first step of this reaction has to include extracting an activated hydrogen from 4 to form an enolate ion which reacts with Eschenmoser’s salt 19 to form intermediate 20.

\[
\text{HCH}_2\text{OR} + \text{H}^+ \rightarrow \text{HCH}_2\text{OR}^+ \quad (4)
\]

In order to check whether or not the alkoxide ion as well as the ion from the lithium diisopropylamide (LDA) can easily extract the α-proton of the propionate ester to form 18. TSs for extracting the α-proton of methyl propionate extractions were calculated. In these calculations, the dimethylamide ion 23 ((CH₃)₂N⁻) is a model for a diisopropylamide ion from LDA. Figure 3 shows structures related to the extraction including solvent molecules and an energy diagram. C-H and N-H lengths in 18c(N⁻, TS, THF) were calculated to be 1.288 and 1.501 Å, respectively. The activation energy turned out to be as small as 1.1 kcal mol⁻¹. The product is more stable than the reactant by 13.6 kcal mol⁻¹. It is concluded that 18c(N⁻, THF) in the reaction mixture exists in a sufficient concentration to proceed the reaction with 19.

The Ea for the methoxide ion was calculated to be 4.5 kcal mol⁻¹ and the product is less stable by 4.0 kcal mol⁻¹ than the reactant. The relative stability between 18c(O⁻, THF) and 4c is different from that of the amide. It is, therefore, considered that the concentration of the ion is too small to react with Eschenmoser’s reagents in the reaction mixtures. The calculated results suggested
that LDA rather than the alkoxide ion should be used in experiments. A minimum energy path calculation showed that formation 20c proceeds without an appreciable barrier.

In the second step, 20 reacts with methyl bromide under an S$_{N}$2 type mechanism, followed by the release of a quaternary ammonium ion from 22 and to form the target in the third step. Figure 2b displays the geometry of 22a(TS), which is higher in energy by 19.7 kcal mol$^{-1}$ than the reactant, 20a + CH$_3$Br. These results suggest that the extraction of the activated hydrogen from 4 easily leads to the formation of the final product.

As this route comes from the TOSP program, we have to consider the reaction conditions for experiments. In the present case, we referred to Poulter and his coworkers in Eq. 6.

We performed experiments for the first step using several conditions and the results are summarized in Scheme 2. No reactions proceeded when we used NaH as the base and the reactant 4a was recovered (Entry 1). Experiments using the anion from t-BuOK recovered 4a and produced benzyl alcohol 24a meaning that the ion was used not to extract the α-hydrogen but to attack the carbonyl carbon of 4a to form a tetrahedral intermediate. The tetrahedral intermediate releases benzyl alcohol. A similar result was obtained for n-BuLi. These experimental results, Entries 2 and 3 suggested that the basicity of t-BuO$^{-}$ as well as the n-butyl ion does not extract the α-hydrogen of 4a. The anions prefer reacting with the carbonyl carbon to extracting the α-hydrogen.

![Scheme 2](image)

It is considered that LDA is a stronger base than H$^+$, alkoxide and alkyl anions. In using the stronger base, we also obtained traces of 25a instead of 20a. This is a product of the Claisen condensation of 4a, i.e., the diisopropylamide ion extracts the α-proton to form 18a while the ion produced reacts not with 19 but 4a. As the theoretical calculations suggest the existence of the route from 19 to 1, further improvements in the reaction conditions may lead to better results for synthesizing 1.

### Route D

The KOSP produces synthetic routes according to the knowledge from a popular database and therefore, provides references that are used for creating routes. In this case, the program cited a paper from Natsugari who examined Eq. 7.

![Reaction Equation](image)

It is considered that solvent effects are deeply related to the reaction in Eq. 8 since the acrylate ion, which attacks benzyl bromide, can make a hydrogen bond with the solvent. The bond weakens the basicity of the ion. At the B3LYP/6-311+G** level of theory, the hydrogen bond energy between the acrylate ion and ethanol was calculated to be 17.5 kcal mol$^{-1}$ and that for acetonitrile to be 14.9 kcal mol$^{-1}$. These results lead to the expectation that the reaction in EtOH requires higher temperatures than that in acetonitrile in order to provide naked ions for the S$_{N}$2 reaction. For the reaction at the same temperature, the yield in EtOH is probably lower than that in CH$_3$CN.

![Equation](image)

Bases such as dicyclohexylamine and NaH are used to extract acid hydrogens from carboxylic acids in organic solvent to form carboxylates. The generated anions attack alkyl halides, and in the present case benzyl bromide, under the S$_{N}$2 mechanism to form 1a as shown in Eq. 8. If the formation of acrylate ions is not difficult, the S$_{N}$2 step should be the rate-determining step. The TS for the reaction is displayed in Figure 2c and the activation barrier was calculated to be only 7.6 kcal mol$^{-1}$. It is considered that this value is small enough to allow the reaction to proceed at room temperature.

Table 1 lists yields of S$_{N}$2 reactions in different solvents. Although the activation energy is very low, we had to heat up reaction mixtures to 50 °C in EtOH for 44 h and the reaction yield was as low as 58% (Entry 3). On the other hand, the yield in CH$_3$CN for 24 h was 88% at room temperature when 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) was used as a base.

The aforementioned is consistent with calculated results concerning hydrogen bond energy strengths between solvents and the acrylate ions.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Base</th>
<th>Solvent</th>
<th>Conditions</th>
<th>Yield / %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NaH</td>
<td>1 : 1</td>
<td>EtOH</td>
<td>rt, 24 h</td>
</tr>
<tr>
<td>2</td>
<td>NaH</td>
<td>1 : 1</td>
<td>EtOH</td>
<td>rt, 92 h</td>
</tr>
<tr>
<td>3</td>
<td>NaH</td>
<td>1 : 1</td>
<td>EtOH</td>
<td>50 °C, 44 h</td>
</tr>
<tr>
<td>4</td>
<td>NaH</td>
<td>2 : 1</td>
<td>EtOH</td>
<td>rt, 64 h</td>
</tr>
<tr>
<td>5</td>
<td>NaH</td>
<td>1 : 2</td>
<td>EtOH</td>
<td>rt, 64 h</td>
</tr>
<tr>
<td>6</td>
<td>DBU</td>
<td>1 : 1</td>
<td>CH$_3$CN</td>
<td>rt, 24 h</td>
</tr>
</tbody>
</table>
Concluding Remarks

In the present study, we investigated synthetic routes for methacrylate esters created by KOSP and TOSP programs from both theoretical and experimental viewpoints. For three routes, which were theoretically confirmed to be feasible routes to synthesize the target compound, we performed experimental syntheses. We came to the following conclusions from theoretical and experimental results:

(1) For Route A, the DFT calculations gave an activation energy for the byproduct lower than that for the target. It is impossible to synthesize 1 using a Favorovskii rearrangement of 2.

(2) Route B and Route D were experimentally confirmed to give the target, benzyl methacrylate. For both routes, we found TSs and activation energies values that are consistent with experimental conditions.

(3) Although we obtained a reasonable reaction path for Route C, the key intermediate 20 of the route was not synthesized. While the DFT calculations proved that both LDA and alkoxide ions are powerful enough to extract α-hydrogens from methyl propionate, we did not succeed in synthesizing the target with Route C. In order to overcome this problem, we have to devise a method for selecting proper solvents for this reaction using theoretical calculations or experimental techniques.

References and Notes


[22] Structures including solvent molecules are expressed by the number with base and solvent. N and O are amide and alkoxide ions, respectively.


計算化学と情報化学の融合による合成経路設計の試み；Ethyl Methacrylate 及び Benzyl Methacrylateへの適用

堀憲次*a, 貞富博喬a, 岡野克彦b, 隅本倫徳a, 宮本敦夫*, 林里織a, 山本豪紀a

標的化合物の合成経路を創出する情報化学的手法を用いた合成経路設計プログラムと、化学反応の機構解明に有力な手段である計算化学との融合による合成経路設計支援システムの開発を目指し、KOSP および TOPS から導出されたメタクリル酸エステル合成経路の理論的・実験的検証を行った。KOSP および TOPS から導出されたメタクリル酸エステルの合成経路のうち、Favorskii 転位、Horner–Emmons 反応、Eschenmoser 試薬を用いる反応、カルボキシラートの S_N2 反応の 4 経路について、B3LYP/6-31G* あるいは、B3LYP/LANL2DZ レベルの密度汎関数理論 (DFT) 計算により、詳細な反応解析を行った。その結果、Favorskii 転位を経る方法では、より安定な副生成物が得られると予想されたが、その他の 3 経路では遷移状態や活性化エネルギーの値から穏やかな条件下で反応が進行すると評価された。実験的検証を行ったところ、Eschenmoser 試薬を用いた反応では、目的の生成物を得ることはできなかったが、Horner–Emmons 反応およびカルボキシラートの S_N2 反応においては、標的化合物が得られ、情報化学的手法により創出された合成経路の評価を、計算化学を用いて行うことの可能性が示された。

キーワード: 計算化学、情報化学、合成経路、メタクリル酸エステル

*a kenji@yamaguchi-u.ac.jp