doi.org/10.26434/chemrxiv.11550057.v1

Discovery of a Difluoroglycine Synthesis Method through Quantum Chemical Calculations

Tsuyoshi Mita, Yu Harabuchi, Satoshi Maeda

Submitted date: 08/01/2020 • Posted date: 09/01/2020

Citation information: Mita, Tsuyoshi; Harabuchi, Yu; Maeda, Satoshi (2020): Discovery of a Difluoroglycine Synthesis Method through Quantum Chemical Calculations. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.11550057.v1

The systematic exploration of synthetic pathways to afford a desired product through quantum chemical calculations remains a considerable challenge. In 2013, Maeda et al. introduced ‘quantum chemistry aided retrosynthetic analysis’ (QCaRA), which uses quantum chemical calculations to search systematically for decomposition paths of the target product and propose a synthesis method. However, until now, no new reactions suggested by QCaRA have been reported to lead to experimental discoveries. Using a difluoroglycine derivative as a target, this study investigated the ability of QCaRA to suggest various synthetic paths to the target without relying on previous data or the knowledge and experience of chemists. Furthermore, experimental verification of the seemingly most promising path led to the discovery of a synthesis method for the difluoroglycine derivative. The extent of the hands-on expertise of chemists required during the verification process was also evaluated. These insights are expected to advance the applicability of QCaRA to the discovery of viable experimental synthetic routes.

Discovery of a Difluoroglycine Synthesis Method through Quantum Chemical Calculations

Tsuyoshi Mita[1,2]*, Yu Harabuchi[1,2,3,4], and Satoshi Maeda[1,2,3,5]*

[1] Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Kita 21 Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0021, Japan

[2] JST, ERATO Maeda Artificial Intelligence for Chemical Reaction Design and Discovery Project, Kita 10 Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan

[3] Department of Chemistry, Faculty of Science, Hokkaido University, Kita 10 Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan

[4] JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

[5] Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS), Tsukuba, Ibaraki 305-0044, Japan

*E-mail: tmita@icredd.hokudai.ac.jp, smaeda@eis.hokudai.ac.jp

Abstract:

The systematic exploration of synthetic pathways to afford a desired product through

quantum chemical calculations remains a considerable challenge. In 2013, Maeda et al. introduced ‘quantum chemistry aided retrosynthetic analysis’ (QCaRA), which uses quantum chemical calculations to search systematically for decomposition paths of the target product and propose a synthesis method. However, until now, no new reactions suggested by QCaRA have been reported to lead to experimental discoveries. Using a difluoroglycine derivative as a target, this study investigated the ability of QCaRA to suggest various synthetic paths to the target without relying on previous data or the knowledge and experience of chemists. Furthermore, experimental verification of the seemingly most promising path led to the discovery of a synthesis method for the difluoroglycine derivative. The extent of the hands-on expertise of chemists required during the verification process was also evaluated. These insights are expected to advance the applicability of QCaRA to the discovery of viable experimental synthetic routes.

Introduction:

Recent developments in computational algorithms and computer technology have allowed quantum chemical calculations to become an essential tool in the field of organic synthesis.[1-5] One of the most important applications of quantum chemical calculations is for the geometry optimization of transition states.[6] In many recent studies, reaction mechanisms have been elucidated by calculating the transition state through quantum chemical

calculations.[1-5] Increasingly, these calculations are being used not only to confirm experimental results but also to design reactions with improved reactivity and selectivity. In general, quantum chemical calculations require an initial guess for the target molecular structure. That is, they can only be used when a reaction mechanism is assumable. Therefore, quantum chemical calculations have never been used to predict a synthesis method for which a similar method was not already available.

Concurrently, extensive efforts have been made to predict synthesis methods based on enormous amounts of experimental data.[7-14] Some of these efforts are based on a concept similar to retrosynthetic analysis, which is traditionally used in organic synthesis.[15] In retrosynthetic analysis, a synthesis method is proposed by dividing the target molecule into fragments and then determining the synthetic equivalents of each fragment. Unfortunately, it is difficult to identify suitable combinations from the existing vast collection of disconnection patterns and synthesis methods. Similar to distinguished organic chemists, who can often make appropriate choices based on their knowledge and experience, there has been a recent focus on developing the ability of machines to make such choices by learning from huge quantities of experimental data.

Despite these efforts, it remains unclear whether it is feasible to propose a synthesis method from scratch using quantum chemical calculations. Towards realizing such predications, we introduced ‘quantum chemistry aided retrosynthetic analysis’ (QCaRA),

which applies quantum chemical calculation to search systematically for the decomposition paths of a target molecule and suggest a synthesis method corresponding to the reverse reaction of the obtained path.[16] This concept is similar to retrosynthetic analysis, as it predicts a reaction path by decomposing the target molecule into fragments and tracing the decomposition pathways backward to generate input molecules. However, QCaRA differs from general retrosynthetic analysis in three ways: (1) the transition state of the decomposition process can be obtained, (2) the effect of the presence of compounds other than the target compound on the decomposition path and its transition state can be predicted, and (3) reactant candidates that are not found in previous data and would typically be considered implausible can be included. Of these, (1) and (2) are both advantageous and disadvantageous. Because the transition state of the decomposition process is also that of the formation process, its energy and structure can guide the design of a synthetic path. Furthermore, executing calculations with a virtual catalyst added to the system enables the selective exploration of paths in which the catalyst effectively lowers the energy barrier. However, because of the vast amount of catalyst options, the result depends on the selection of a suitable catalyst, which is a serious drawback.

Previously, the prediction capability of QCaRA has been considered theoretical, as no new reactions have been discovered experimentally based on the synthetic paths proposed by QCaRA. Maeda et al. introduced QCaRA based on the results of an extensive exploration of

the decomposition paths of glycine. Prior to this analysis, they had reported that glycine could be obtained from acetolactone and ammonia or ammonium ylide and CO_2 as reactant pairs.[17,18] Separately, Mita et al. had developed a synthesis method involving the carboxylation of ammonium ylide with CO_2.[19] Although the findings of Maeda et al. and Mita et al. were separate outcomes in different fields, this case can be considered to demonstrate the possibility of experimentally verifying a prediction made through QCaRA.

Herein, we report the first successful discovery of a synthesis method based on a theoretical path suggested by QCaRA. In this study, to demonstrate the power of QCaRA, we chose a target molecule for which no facile synthesis methods exist, i.e. α,α-difluoroglycine. Replacement of a hydrogen atom with a fluoride atom in bioactive molecules such as α-amino acids is a promising strategy for enhancing biological activity and bioavailability via increased lipophilicity and improved stability against enzymatic degradation. It should be emphasized that the theoretical path that led the discovery of the synthesis method originated solely from quantum chemical calculations without the input of any previous experimental data or human knowledge. Nevertheless, some steps in the experimental verification process demanded human knowledge and experience. Thus, this paper discusses the contribution of quantum chemical calculations to the prediction of synthetic methods and when human knowledge is required.

Results:

Prediction of synthetic paths using QCaRA

Over the last several years, the development of various automated reaction path search methods has shown that many reaction paths can be predicted.[20-22] If previously unknown chemical transformations were to be predicted and such novel chemical transformations achieved experimentally, a series of new chemical reactions could be discovered. This study used QCaRA, proposed by Maeda et al. in 2013, to carry out systematic searches for the decomposition paths of α,α-difluoroglycine as a target molecule and suggest a synthetic path that used the resulting decomposition products or their equivalents as reactants.[16]

The technical difficulty of QCaRA is that the search target is a reaction path that does not actually proceed. Conventional automated reaction path searches can be performed by assuming thermodynamics or kinetics. In other words, it is not necessary to search for a path leading to the thermodynamically unstable structure or for a path with a high energy barrier. In QCaRA, by contrast, the decomposition product that is energetically less stable than the target molecule is often more favourable in the forward synthetic path (exothermic). Thermodynamically very unstable decomposition products should also be searched for because their equivalents could be candidate reactants for novel reactions. Therefore, in this study, an artificial force induced reaction (AFIR) method,[23,24] which has been proven to provide an exhaustive search, including for unstable decomposition products,[25] was adopted

as an automated reaction path search method for QCaRA.

The AFIR method induces structural changes by applying a virtual artificial force between molecules or fragments within a molecule. The systematic repetition of this process makes it possible to calculate a path for transforming a given reactant into an unknown product. It also allows the prediction of an unknown reaction by analysing the network of obtained reaction paths. The further details of this method are beyond the scope of this paper.[23,24] In this study, the AFIR method was only applied to structures with the same bonding pattern as the input molecule and the decomposition path was comprehensively searched.[26] After prospective decomposition products were selected from the obtained candidates, the automated reaction path search was also performed using their equivalents as starting materials, where the option to select a path kinetically was used.[27] Finally, the predicted synthetic path was experimentally verified. The calculations and experiments are detailed in the Supplementary Information.

Prediction of synthetic methods for α,α-difluoroglycine

This study focused on the synthesis of α,α-difluoroglycine ($NH_2CF_2CO_2H$), for which an efficient synthesis method has not yet been established. However, obtaining such a synthesis protocol would be highly desirable due to the expected usefulness. Therefore, the difluoroglycine derivative was used as an input when executing the automated reaction path

search. A number of reactant candidates were obtained through a constrained search, in which the AFIR method was only applied to the equilibrium structures with the same bonding pattern as the input structure, i.e. conformers of difluoroglycine.

Of the obtained reactant candidates, the 30 with the lowest energies are shown in Figure 1. These reactant candidates correspond to decomposed species obtained by the systematic exploration of the dissociation pathways of difluoroglycine. The most stable reactant pair is R1, which consists of difluoromethylamine and CO_2. The set with the second lowest energy (R2) consists of fluoroimine, CO_2, and hydrogen fluoride. Although most species observed in these 30 reactant candidate sets are stable molecules in which all atoms fulfil the octet rule, there are also some short-lived intermediates like carbene. The energy levels of the reactant candidates are depicted in Figure 2 together with the reaction path network on which these species were predicted. Notably, the 4 sets with the lowest energies, i.e. R1–R4, cannot be reactant candidates because the reactions from them to the target molecule are endothermic. Among the other 26 reaction candidates, the path from R26 (CF_2, NH_3, and CO_2), highlighted in Figures 1 and 2, was selected considering the energy barrier for the formation of difluoroglycine and the availability of the reactants. As seen in the reaction profile for this path (Figure 3), the energy barrier along the path from R26 is reasonably small, and the three simple components, i.e. CF_2, NH_3, and CO_2, convergently react to give difluoroglycine in one step.

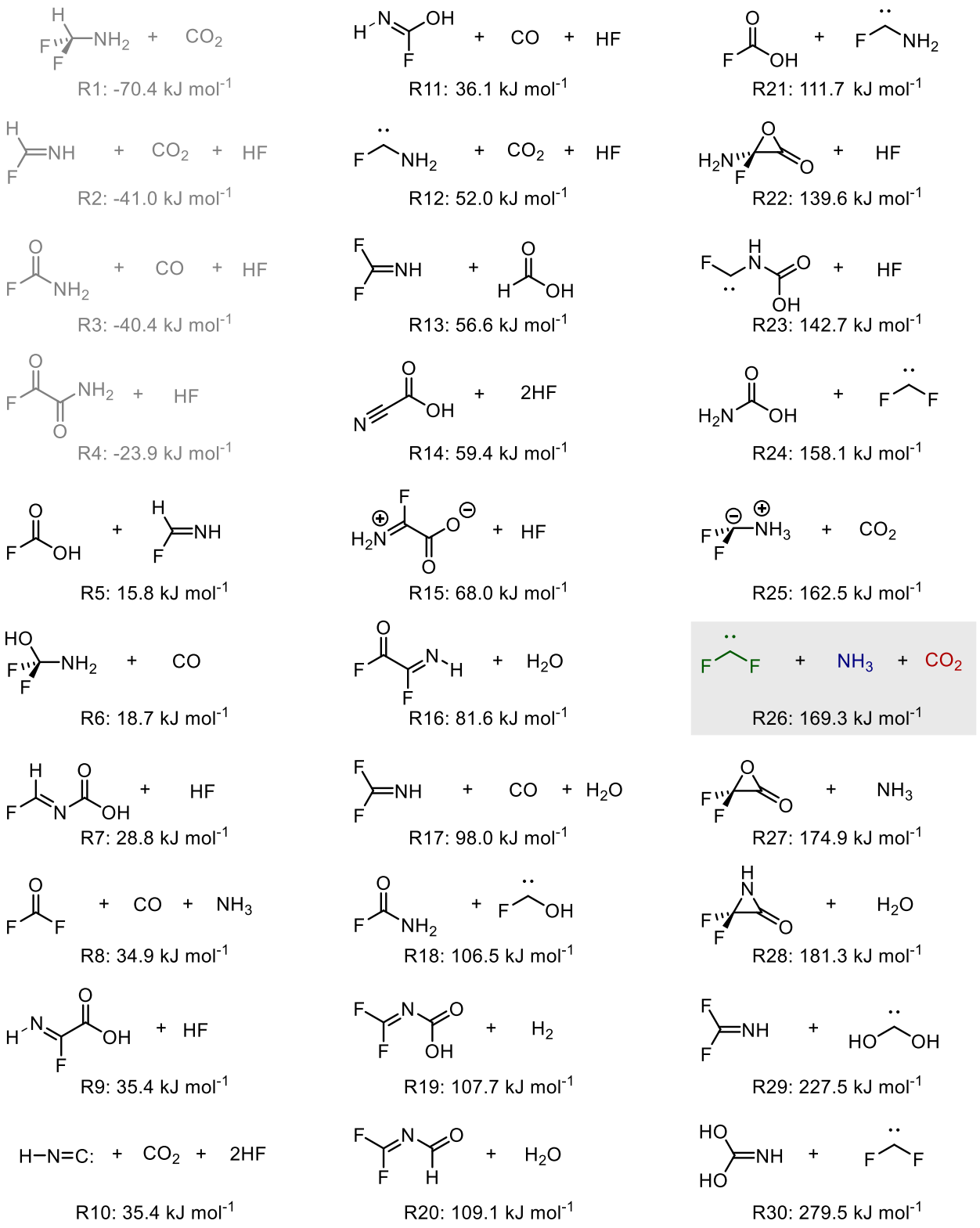


Figure 1. Reactant candidates predicted by the automated reaction path search.

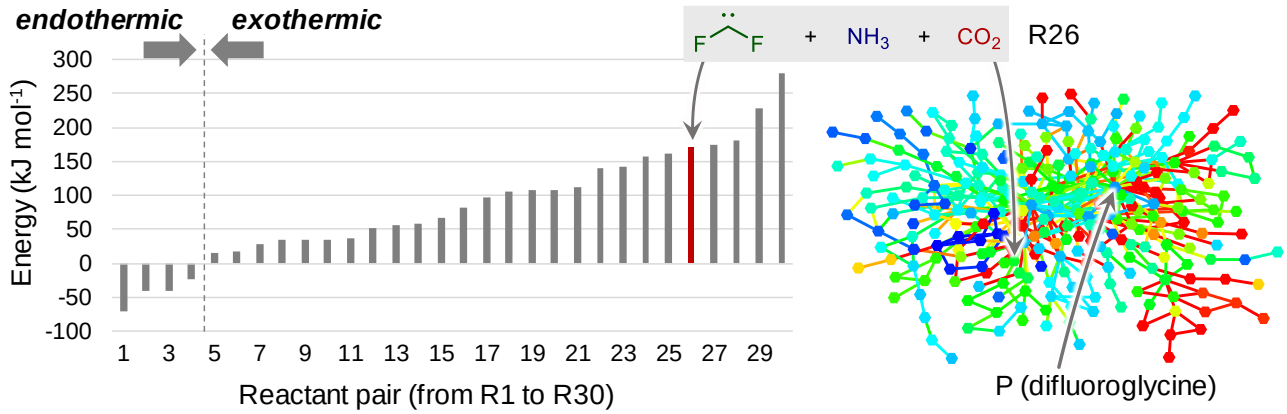


Figure 2. Histogram of the energy levels of the 30 reactant candidates (R1–R30, listed in Figure 1) (left) and the reaction path network obtained through the automated reaction path search (right). In the reaction path network, each node and edge is colour-coded according to the energy of the corresponding equilibrium structure and transition state (blue: −70.4 kJ mol^{-1}, aqua: 54.5 kJ mol^{-1}, green: 179.5 kJ mol^{-1}, yellow: 304.5 kJ mol^{-1}, and red: 429.5 kJ mol^{-1}).

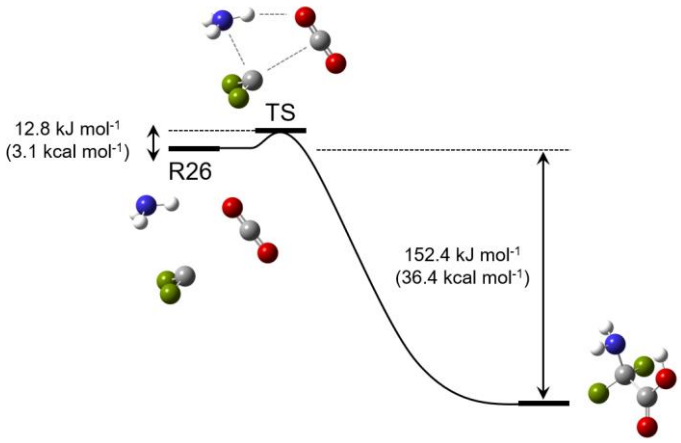


Figure 3. Reaction path for R26, which was selected for experimental demonstration. The H, C, N, O, and F atoms are coloured white, grey, blue, red, and green, respectively.

Among the components of R26, NH_3 and CO_2 are both commercially available. In particular, CO_2 is an abundant, inexpensive, and non-toxic chemical utilized in various organic transformations as a C1 unit.[28] Furthermore, difluorocarbene can be generated in

situ in several ways.[29] Among them, a method using CF_3^- as a precursor, which is generated from Me_3SiCF_3 (Ruppert–Prakash reagent) and $Ph_3SiF_2{\cdot}NBu_4$ (TBAT), was initially selected owing to the reasonable accessibility of both reagents.[30] To evaluate the validity of this synthesis method, the automated reaction path search was executed using CF_3^- + NH_3 + CO_2 as reactants, with the constrained search option applying the AFIR method only to equilibrium structures that the system can reach at a reaction temperature of 300 K in a reaction time of 1 h. The resulting reaction path networks are illustrated in Figures 4a and 4b, which represent identical networks colour-coded by (a) the energy of the structure to which each node corresponds and (b) the calculated yield of the corresponding structure. The calculated yield was obtained by simulating the propagation of the population from the initial structure using the rate constant matrix reaction method under a reaction temperature of 300 K and a reaction time of 1 h. Figure 4a shows that the target product, difluoroglycine, is obtained on the network, while Figure 4b indicates that the calculated yield of difluoroglycine is almost zero because the equilibrium between CF_3^- and CF_2 + F^- favours the former. As a result, $CF_3CO_2^-$, in which CF_3^- is directly bound to CO_2, was obtained as the main product (99.8%).

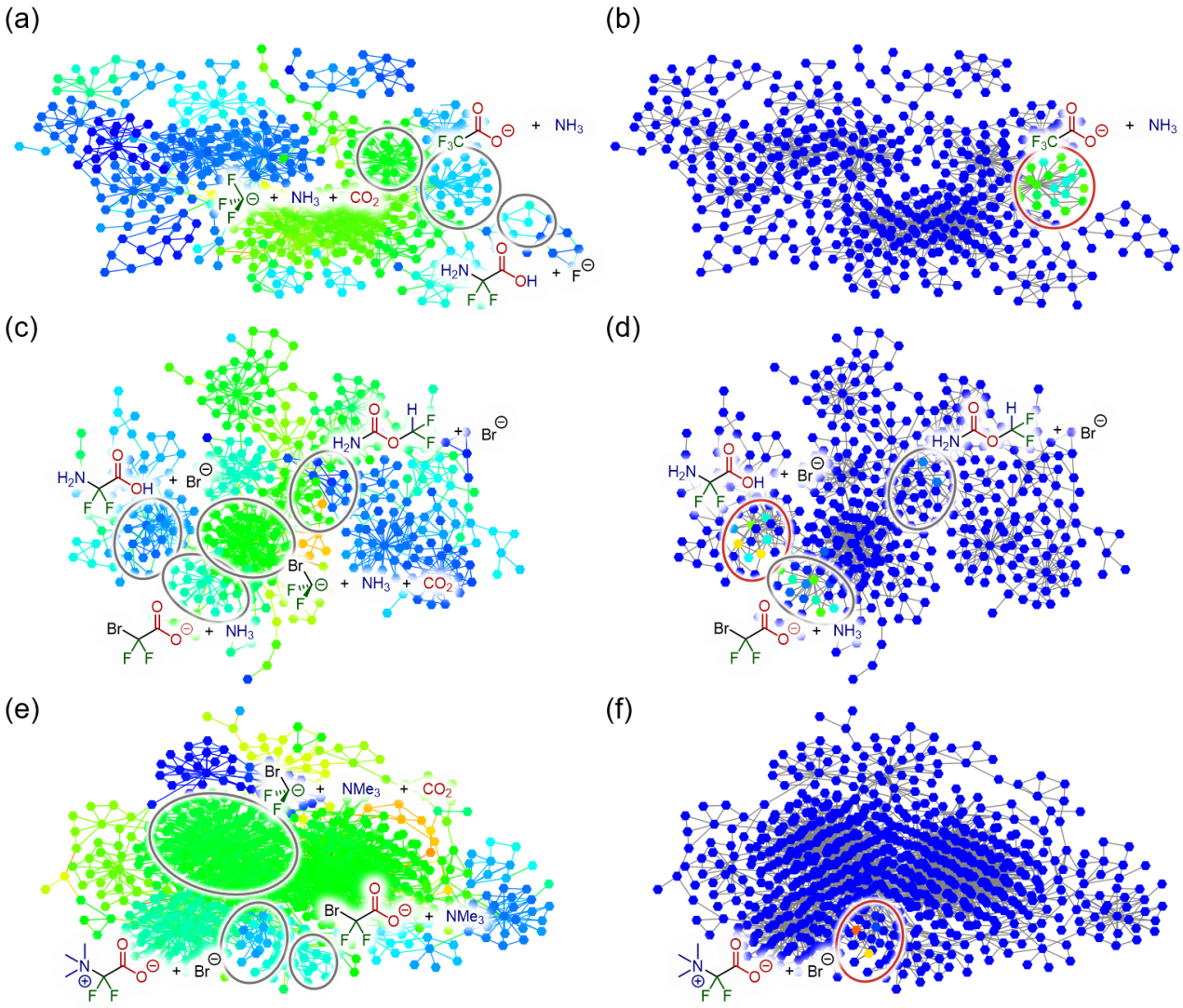


Figure 4. Reaction path networks corresponding to (a, b) CF_3^- + NH_3 + CO_2, (c, d) CF_2Br^- + NH_3 + CO_2, and (e, f) CF_2Br^- + NMe_3 + CO_2. Each node and edge in the reaction path networks of (a), (c), and (e) is colour-coded by the energy (relative to the most stable structure) of the corresponding equilibrium structure and transition state (blue: 0.0 kJ mol^{-1}, aqua: 125.0 kJ mol^{-1}, green: 250.0 kJ mol^{-1}, yellow: 375.0 kJ mol^{-1}, and red: 500.0 kJ mol^{-1}). Each node in the reaction path networks of (b), (d), and (f) is colour-coded by the calculated yield of the corresponding equilibrium structure at a reaction temperature of 300 K and a reaction time of 1 h (blue: 0.0%, aqua: 0.8%, green: 4.0%, yellow: 20.0%, and red: 100.0%).

As the addition of CF_3^- to CO_2 was faster than α-elimination from CF_3^-, to enhance α-elimination, Me_3SiCF_2Br was then chosen as the difluorocarbene precursor. The improved

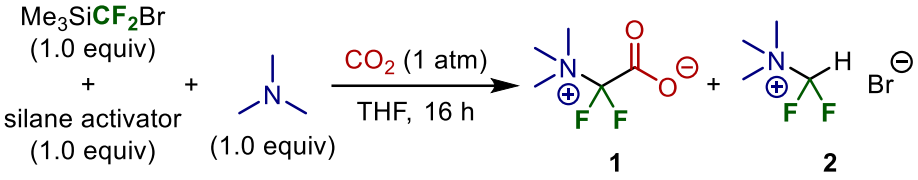
leaving ability of bromide accelerates α-elimination from the corresponding CF_2Br^-, which can be generated in the presence of an appropriate silane activator.[31] To examine the validity of this method, the reaction path networks were obtained with $CF_2Br^- + NH_3 + CO_2$ as reactants (Figures 4c and 4d). The usage of CF_2Br^- was revealed to shift the equilibrium between CF_2Br^- and $CF_2 + Br^-$ towards the latter; thus, difluoroglycine was obtained. However, as shown in Figure 4d, the generation of $CF_2BrCO_2^-$ (29.3%) competes with the process for forming the target product (69.6%). Furthermore, it was suggested that $NH_2CO_2CHF_2$ was obtained as a minor by-product (0.8%).

To further improve the yield of difluoroglycine, the addition of the amine to difluorocarbene should be accelerated. It is also necessary to reduce proton transfer from the amine to CF_2 in order to suppress the formation of $NH_2CO_2CHF_2$. Therefore, to meet both of these requirements simultaneously, a tertiary amine was selected. To test the validity of this hypothesis, the reaction path networks with $CF_2Br^- + NMe_3 + CO_2$ as reactants were obtained, as illustrated in Figures 4e and 4f. These results verify that the desired difluoroglycine derivative was selectively obtained when NMe_3 was used as a reactant. In addition, the calculated yield of 99.98% could be further increased to 99.99% by lowering the reaction temperature to 250 K and reached almost 100% when the temperature was further lowered to 200 K.

Experimental verification of the predicted synthetic method

As a calculated yield of 100% was achieved with CF_2Br^- + NMe_3 + CO_2 as reactants, we chose to verify the validity of this synthetic path for the difluoroglycine derivative experimentally. We first examined several silane activators for the generation of the difluorocarbene in situ from Me_3SiCF_2Br in THF under an atmosphere of 1 atm CO_2 at room temperature (~300 K) (Table 1). When CsF was employed as the silane activator, target difluoroglycine derivative **1** was obtained in 18% yield together with compound **2**, which was derived from protonation of the ammonium ylide intermediate, in 27% yield. Although $F{\cdot}NBu_4$ (TBAF) was not a suitable activator, the use of TBAT dramatically improved the yield of **1** to 81% and gave **2** in only 7% yield. Subsequently, the reaction was carried out at a lower temperature (−40 °C = 233 K) for 16 h and the generated precipitates were isolated by filtration. It was found that the obtained solids, as a mixture with $Br{\cdot}NBu_4$, contained **1** and **2** in 96% and 2% yields, respectively. This result, with **1** obtained in high yield, is consistent with our prediction based on quantum chemical calculations. In addition, we conducted the reaction of Me_3SiCF_3 (1 equiv), TBAT (1 equiv), and NMe_3 (1 equiv) under a CO_2 atmosphere (1 atm) in THF at room temperature for 16 h. This reaction gave trifluoroacetate ($CF_3CO_2{\cdot}NBu_4$) in 96% yield, which is also consistent with the theoretical prediction that CF_3^- is consumed quickly by reacting with CO_2 rather than providing the difluorocarbene.

Table 1. Condition Screening for Silane Activators



1 **2**

entry	silane activator	temp (°C)	yield of **1** (%)[a]	yield of **2** (%)[a]
1	CsF	r.t.	10	27
2	F•NBu_4 (TBAF)	r.t.	trace	66
3	Ph_3SiF_2•NBu_4 (TBAT)	r.t.	81	7
4	Ph_3SiF_2•NBu_4	-40	**96**	2

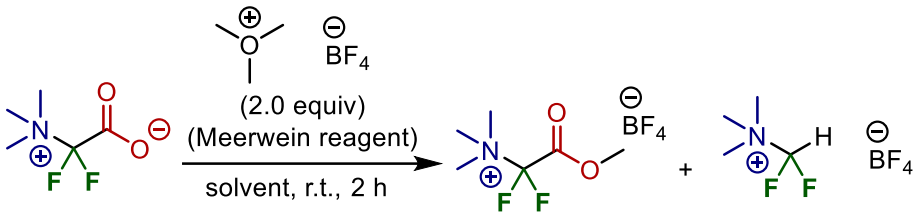
[a]Yields were determined by ^{1}H NMR analysis in DMSO-d_6 using pyridine as an internal standard.

Although the reaction occurred in THF, product **1** was obtained as a precipitate that was insoluble in THF. Thus, the precipitate should be dissolved in another solvent for further purification to remove Br·NBu_4. However, the dissolution of salt **1** in solvents such as H_2O, MeOH, CH_3CN, and DMF promoted to varying extents a decarboxylation–protonation process that generated **2**. For example, when the obtained product mixture (**1**:**2** = 98:2) was dissolved in various solvents at room temperature for 1 h, the ratio of **1**:**2** changed to 87:13 (H_2O), 89:11 (MeOH), 62:38 (CH_3CN), and 50:50 (DMF). This result indicates that it might be difficult to achieve the purification of **1** at this stage.

Therefore, esterification was investigated as a method of isolating and purifying the difluoroglycine derivative without decarboxylation. First, MeI was employed as a methylating reagent. However, an established condition using MeI in DMF or an excess amount of MeI without any solvent (neat conditions) did not afford the corresponding methyl carboxylate.

Instead, the undesired decarboxylation–protonation process proceeded to some extent. Thus, esterification using highly electrophilic $BF_4 \cdot OMe_3$ (Meerwein reagent)[32] was investigated. However, as the esterification process also requires redissolution in a highly polar solvent in which salt **1** is soluble, there was some concern that the yield might be decreased. First, CH_2Cl_2, which is a typical solvent for esterification using the Meerwein reagent, was used (Table 2). Although target methyl ester **3** was obtained in 18% yield, protonated compound **4** was obtained as a by-product in 73% yield. In contrast, when esterification was performed under solvent-free ball-milling conditions, the yield of **3** was increased to 48%. Furthermore, although AcOEt was not a suitable solvent, the use of acetone as a solvent improved the yield of **3** to 81%.

Table 2. Condition Screening for Esterification



1 (**1**:**2** = 98:2) **3** **4**

entry	solvent	yield of **3** (%)[a]	yield of **4** (%)[a]	recovery of **1** (%)[a]
1	CH_2Cl_2	18	73	8
2	- (ball milling)	48	12	36
3	AcOEt	4	80	15
4	acetone	**81**	16	-

[a]Yields were determined by ^{1}H NMR analysis in DMSO-d_6 using pyridine as an internal standard.

Subsequently, we conducted preparative-scale synthesis using 1 mmol of each reactant under the optimized reaction conditions (Figure 5). After **1** was obtained as a mixture with $Br{\cdot}NBu_4$, the obtained solids were treated with $BF_4{\cdot}OMe_3$ in acetone. The resulting product (methyl ester **3**) was washed with MeOH to remove $Br{\cdot}NBu_4$, unreacted $BF_4{\cdot}OMe_3$, and protonated compound **4**, affording **3** in 80% yield (205 mg). The structure was confirmed by X-ray crystallography after recrystallization from MeOH (CCDC 1971834). This method is an elegant demonstration of the synthesis of a difluoroglycine derivative from three simple compounds through multi-component assembly under silica-gel-column-chromatography-free conditions. In most reported examples,[29] an excess amount of difluorocarbene precursor (>2 equiv) is necessary, probably due to undesirable carbene dimerization. In contrast, our new protocol achieves a high yield with just 1 equiv of Me_3SiCF_2Br, which emphasizes the practicality of this synthesis. Other substrate candidates predicted by the quantum chemical calculations are now being investigated and the results will be reported in due course.

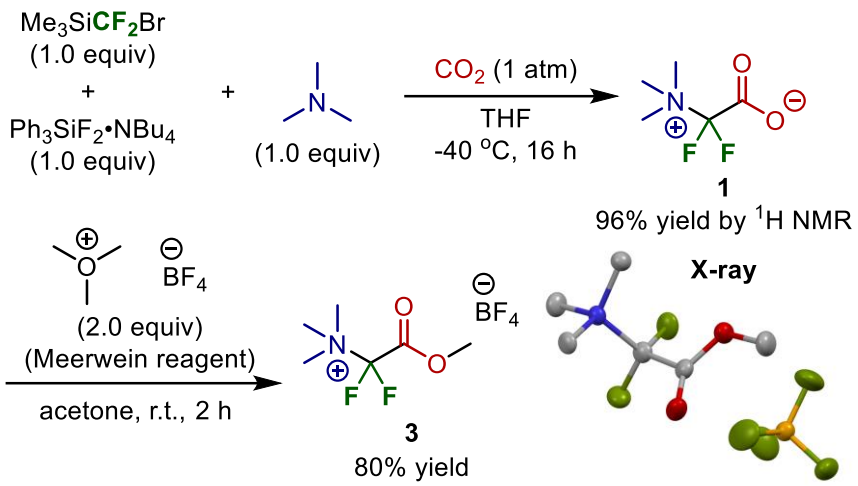


Figure 5. Preparative-scale synthesis of a difluoroglycine derivative without silica-gel column chromatography.

Discussion:

We would like to highlight that this synthesis of a difluoroglycine derivative is the first experimental demonstration of a reaction suggested by QCaRA. Moreover, the process by which QCaRA provided predictions, i.e. the process for obtaining reactant candidates (Figure 1) after deciding the target, did not rely on the experience or intuition of chemists nor was any previous experimental data used in this process. Although the developer version of the GRRM program was used in this study, the same calculation can be performed using the GRRM17 program.[26] It is noted that a chemist's knowledge was indispensable in selecting difluoroglycine as the target and choosing R26 from the many candidates predicted by QCaRA. Automating these two processes remains a future challenge.

Furthermore, the experience and intuition of chemists also played an important role in

validating the predicted synthetic path. First, because the proposed reactant candidates included a molecule (difluorocarbene) that was not commercially available, it was generated in situ using a previously reported method. The calculations to determine what products were actually produced and in what proportion suggested that the yield of the target compound could be improved by changing the type of amine. The use of a tertiary amine based on this prediction was also decided by chemists according to their experience.

The process that demanded the greatest input from chemists was the isolation of products. After determining the optimal reagents and reaction conditions, the product was precipitated as a mixture with $Br{\cdot}NBu_4$; however, the target difluoroglycine derivative was not sufficiently stable during subsequent isolation processes. Because the calculated yield was approximately 100%, the decision was taken that isolation should be performed after methyl esterification. After examining several solvents for this process, acetone was found to be the most suitable. In fact, the trial and error required for this step was the most time-consuming process in this study.

The selection of both THF as the reaction solvent and the calculation method were also responsible for the success of this demonstration. THF was chosen as the reaction solvent because it is commonly used in CO_2 fixation and carbene insertion chemistry. However, if a different solvent had been used, competition with decarboxylation might have prevented isolation of the target product. For the computational techniques, it is possible that the

choice of an unsuitable DFT functional or basis function could lead to incorrect predictions. It is undeniable that the experience and intuition of chemists, or even luck, contributed to appropriate choices being made.

Conclusions:

In this study, we applied QCaRA to a difluoroglycine derivative for which no efficient synthesis method has been established. Out of the 30 synthetic paths proposed for difluoroglycine, those involving a difluorocarbene, an amine, and CO_2 were selected for further investigation. The effects of the difluorocarbene generation method (CF_3^- vs CF_2Br^-) and the selected amine (NH_3 vs NMe_3) on the yield were then verified by calculation to determine a set of reactants that gave a calculated yield of almost 100%. Subsequently, experiments were conducted to verify the predicted synthetic path.

After the synthetic target was chosen, the prediction of potential synthetic paths by QCaRA did not require the experience and intuition of chemists or previous experimental data. However, the experience of chemists and trial-and-error experiments were necessary during the verification process. In this example, this hands-on expertise of chemists was needed to select a suitable path from among those predicted by QCaRA, to propose a difluorocarbene production method, to select an appropriate amine, and to isolate the product. In particular, experimental trial and error was key for the purification of the product.

This study demonstrated the effectiveness of QCaRA for predicting new synthesis methods. However, if the procedures used in this study were applied to a more complex molecule, the computational cost would be huge. In particular, for a catalytic reaction, QCaRA would need to be performed on the system to which the catalyst is added, which would further increase the computational cost. Moreover, to evaluate different catalysts systematically, QCaRA would need to be repeated while considering various metal/ligand combinations. In future, we hope to apply QCaRA to more complex systems by improving the calculation procedures, making a database of the results, and so forth.

Experimental section:

In an oven-dried round-bottom flask was placed TBAT (539.9 mg, 1.0 mmol, 1 equiv). The flask was evacuated and backfilled with CO_2 (3 times) followed by the addition of THF (10 mL). The mixture was stirred at room temperature until TBAT was completely dissolved. After adding trimethylamine (2 M in THF, 500 μL, 1.0 mmol, 1 equiv) at 0 °C, the solution was cooled to −40 °C, and then Me_3SiCF_2Br (156 μL, 1.0 mmol, 1 equiv) was added dropwise. After the resulting slurry was stirred at −40 °C for 20 h, the precipitate was isolated by filtration, washed with hexane, and dried under vacuum. The approximate yield of carboxylate **1** was 96%, as determined by ^1H NMR spectroscopy using pyridine as an internal standard. The solids were then treated with the Meerwein reagent (295.8 mg, 2.0

mmol, 2 equiv) in acetone (1.5 mL) for 2 h. After the solvent was removed under vacuum, the resulting solids were washed with a small amount of MeOH to afford methyl carboxylate **3** after filtration (1st crop: 176.6 mg, 2nd and 3rd crops: 28.2 mg, 0.803 mmol, 80% yield). White solids; IR (ATR): 3068, 2977, 1787, 1483, 1343 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6) δ: 4.05 (s, 3H), 3.40 (s, 9H) ppm; ^{13}C NMR (100 MHz, DMSO-d_6) δ: 155.7 (J_{CF} = 31.6 Hz), 112.5 (J_{CF} = 282.7 Hz), 56.5, 49.4 ppm; ^{19}F NMR (376 MHz, DMSO-d_6) δ: −105.4 (CF_2), −151.4 (BF_4) ppm (internal reference: CF_3CO_2H in DMSO-d_6 = −78.5 ppm); HRMS (ESI) *m*/*z* calcd. for $C_6H_{12}F_2NO_2^+$ [M-BF_4^-]$^+$: 168.0831, found: 168.0833; calcd. for BF_4^- [M-$C_6H_{12}F_2NO_2^+$]: 87.0035, found: 87.0030.

Computational analysis:

All the calculations were performed using the developer version of the GRRM program combined with the Gaussian 16 program.[33] For all the calculations, the ωB97X-D functional and 6-31+G* basis functions were used and the Grid=FineGrid option was adopted. The Gibbs free energy values at 300 K and 1 atm were estimated by assuming ideal-gas, rigid-rotor, and harmonic vibrational models, where all the harmonic frequencies smaller than 50 cm^{-1} were set as 50 cm^{-1} as suggested in the literature.[34] The QCaRA calculation was performed under gas-phase conditions, whereas the other calculations were performed in THF, as modelled by the CPCM method, because THF is often chosen as an experimental

solvent for reactions involving CO_2. Kinetic simulations were performed using the reaction path networks shown in Figure 4, where the initial populations given to the reactants were propagated using the rate constant matrix reaction method at a reaction temperature of 300 K and a reaction time of 1 h.[27] Further details on the reaction path searches and kinetic simulations are available in the Supplementary Information.

References:

1) Houk, K. N. & Cheong, P. H.-Y. Computational prediction of small-molecule catalysts. *Nature* **455**, 309-313 (2008).
2) Thiel, W. Computational catalysis–past, present, and future. *Angew. Chem., Int. Ed.* **53**, 8605-8613 (2014).
3) Sameera, W. M. C., Maeda, S. & Morokuma, K. Computational catalysis using the artificial force induced reaction method. *Acc. Chem. Res.* **49**, 763-773 (2016).
4) Houk, K. N. & Liu, F. Holy grails for computational organic chemistry and biochemistry. *Acc. Chem. Res.* **50**, 539-543 (2017).
5) Ahn, S. et al. Design and optimization of catalysts based on mechanistic insights derived from quantum chemical reaction modeling. *Chem. Rev.* **119**, 6509-6560 (2019).
6) Schlegel, H. B. Geometry optimization. *WIREs Comput. Mol. Sci.* **1**, 790-809 (2011).
7) Corey, E. J., Long, A. K. & Rubenstein, S. D. Computer-assisted analysis in organic

synthesis. *Science* **228**, 408-418 (1985).

8) Satoh, H. & Funatsu, K. SOPHIA, a knowledge base-guided reaction prediction system-utilization of a knowledge base derived from a reaction database. *J. Chem. Inf. Model.* **35**, 34-44 (1995).

9) Lin, A. I. et al. Automatized assessment of protective group reactivity: a step toward big reaction data analysis. *J. Chem. Inf. Model.* **56**, 2140-2148 (2016).

10) Szymkuć, S. et al. Computer-assisted synthetic planning: the end of the beginning. *Angew. Chem., Int. Ed.* **55**, 5904-5937 (2016).

11) Wei, J. N., Duvenaud, D. & Aspuru-Guzik, A. Neural networks for the prediction of organic chemistry reactions. *ACS Cent. Sci.* **2**, 725-732 (2016).

12) Coley, C. W., Green, W. H. & Jensen, K. F. Machine learning in computer-aided synthesis planning. *Acc. Chem. Res.* **51**, 1281-1289 (2018).

13) Schwaller, P. et al. “Found in translation”: predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models. *Chem. Sci.* **9**, 6091-6098 (2018).

14) Segler, M. H., Preuss, M. & Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555**, 604-610 (2018).

15) Coley, E. J. The Logic of chemical synthesis: multistep synthesis of complex carbogenic molecules (Nobel lecture). *Angew. Chem., Int. Ed.* **30**, 455-465 (1991).

16) Maeda, S., Ohno, K. & Morokuma, K. Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy by the ADDF and AFIR methods. *Phys. Chem. Chem. Phys.* **15**, 3683-3701 (2013).

17) Maeda, S. & Ohno, K. Ab initio studies on synthetic routes of glycine from simple molecules via ammonolysis of acetolactone: applications of the scaled hypersphere search method. *Chem. Lett.* **33**, 1372-1373 (2004).

18) Maeda, S. & Ohno, K. No activation barrier synthetic route of glycine from simple molecules (NH_3, CH_2, and CO_2) via carboxylation of ammonium ylide: a theoretical study by the scaled hypersphere search method. *Chem. Phys. Lett.* **398**, 240-244 (2004).

19) Mita, T., Sugawara, M. & Sato, Y. One-pot synthesis of α-amino acids through carboxylation of ammonium ylides with CO_2 followed by alkyl migration. *J. Org. Chem.* **81**, 5236-5243 (2016).

20) Zimmerman, P. M. Methods for exploring reaction space in molecular systems. *WIREs Comput. Mol. Sci.* **8**, e1354 (2018).

21) Simm, G. N., Vaucher, A. C. & Reiher, M. Exploration of reaction pathways and chemical transformation networks. *J. Phys. Chem. A* **123**, 385-399 (2019).

22) Sumiya, Y. & Maeda, S. Paths of chemical reactions and their networks: from geometry optimization to automated search and systematic analysis, in Chemical Modelling: Volume 15 (Specialist Periodical Reports) ed. by Springborg, M. & Joswig, J.-O. Royal

Society of Chemistry, pp. 28-69 (2020).

23) Maeda, S. & Morokuma, K. Finding reaction pathways of type A + B → X: toward systematic prediction of reaction mechanisms. *J. Chem. Theory Comput.* **7**, 2335-2345 (2011).

24) Maeda, S. et al. Artificial force induced reaction (AFIR) method for exploring quantum chemical potential energy surfaces. *Chem. Rec.* **16**, 2232-2248 (2016).

25) Maeda, S. & Harabuchi, Y. On benchmarking of automated methods for performing exhaustive reaction path search. *J. Chem. Theory Comput.* **15**, 2111-2115 (2019).

26) Maeda, S. et al. Implementation and performance of the artificial force induced reaction method in the GRRM17 program. *J. Comput. Chem.* **39**, 233-251 (2018).

27) Sumiya, Y. & Maeda, S. A reaction path network for Wöhler's urea synthesis. *Chem. Lett.* **48**, 47-50 (2019).

28) Mita, T. & Sato, Y. Syntheses of α-amino acids by using CO_2 as a C1 source. *Chem. Asian J.* **14**, 2038-2047 (2019).

29) Ni, C. & Hu, J. Recent advances in the synthetic application of difluorocarbene. *Synthesis*, **46**, 842-863 (2014).

30) Wang, F. et al. Synthesis of *gem*-difluorinated cyclopropanes and cyclopropenes: trifluoromethyltrimethylsilane as a difluorocarbene source. *Angew. Chem., Int. Ed.* **50**, 7153-7157 (2011).

31) Li, L., Wang, F., Ni, C. & Hu, J. Synthesis of *gem*-difluorocyclopropa(e)nes and O-, S-, N-, and P-difluoromethylated compounds with $TMSCF_2Br$. *Angew. Chem., Int. Ed.* **52**, 12390-12394 (2013).

32) Raber, D. J., Gariano Jr., P., Brod, A. O. & Gariano, A. Esterification of carboxylic acids with trialkyloxonium salts. *J. Org. Chem.* **44**, 1149-1154 (1979).

33) Gaussian 16, Revision B.01, Frisch, M. J. et al. Gaussian, Inc., Wallingford CT, 2016.

34) Ribeiro, R. F., Marenich, A. V., Cramer, C. J., Truhlar, D. G. Use of solution-phase vibrational frequencies in continuum models for the free energy of solvation. *J. Phys. Chem. B* **115**, 14556-14562 (2011).

Acknowledgements:

This work was financially supported by Grant-in-Aid for Scientific Research (C) (No. 18K05096), JST-CREST (No. JPMJCR14L5), JST-PRESTO (No. JPMJPR16N8), JST-ERATO (No. JPMJER1903), and JSPS-WPI. T.M. thanks the Astellas Foundation for Research on Metabolic Disorders for financial support. We thank Ms. Takako Homma for editing a draft of this manuscript.

Author contributions:

T.M. and S.M. designed the research. T.M. performed the synthetic experiments. Y.H. and

S.M. performed the computational studies. T.M., Y.H., and S.M. prepared the manuscript.

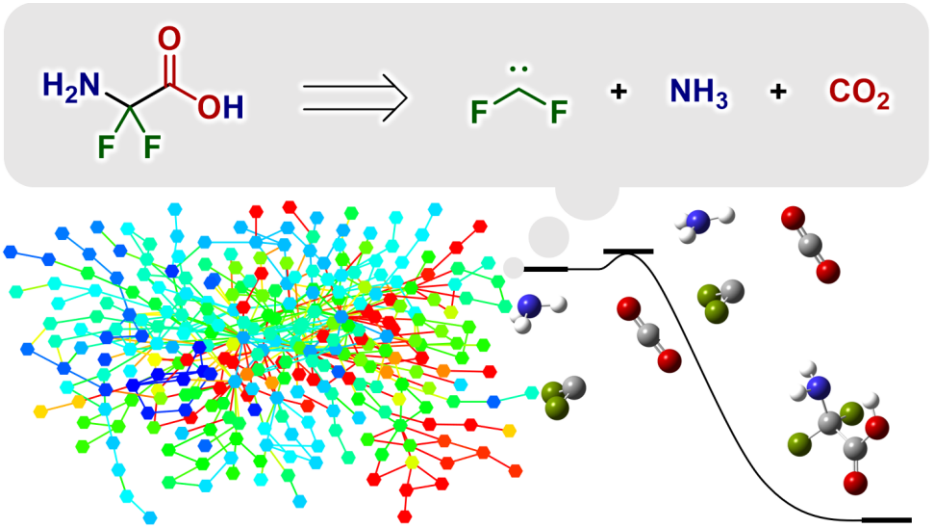
Competing interests:

The authors declare no competing interests.

Additional information:

Supplementary information is available for this paper at https://XXXXX.

Graphical abstract:



Supplementary Information

Discovery of a Difluoroglycine Synthetic Method through Quantum Chemical Calculations

Tsuyoshi Mita,[1,2]* Yu Harabuchi,[1,2,3,4] and Satoshi Maeda[1,2,3,5]*

[1] *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Kita 21, Nishi 10, Kita-ku, Sapporo 001-0021, Japan*

[2] *JST, ERATO Maeda Artificial Intelligence for Chemical Reaction Design and Discovery Project, Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan*

[3] *Department of Chemistry, Faculty of Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan*

[4] *JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan*

[5] *Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS), Tsukuba 305-0044, Japan*

tmita@icredd.hokudai.ac.jp

smaeda@eis.hokudai.ac.jp

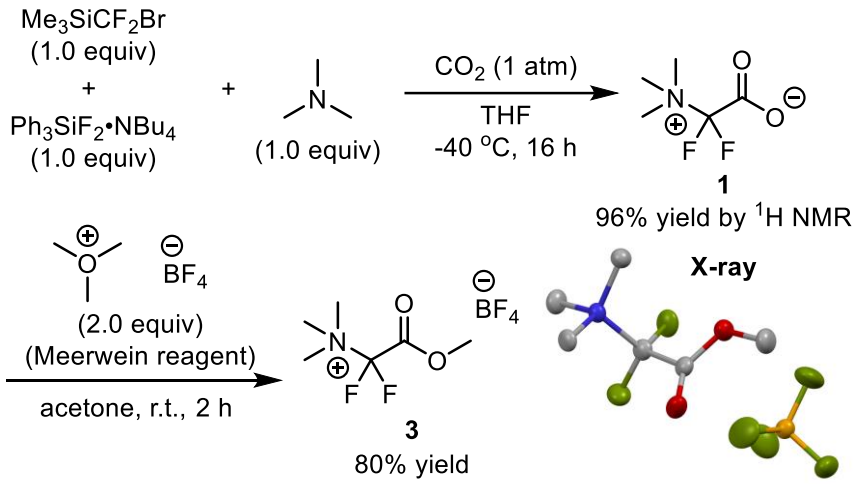
Table of Contents

(A)	General	S3
(B)	The Synthesis of Difluoroglycine Derivative	S3
(C)	Data for Single Crystal X-ray Structural Analysis	S7
(D)	Computational Details	S7
(E)	Predicted Reactant Candidates	S9

(A) General

All manipulations were carried out under an atmosphere of nitrogen unless otherwise noted. Infrared (IR) spectra were recorded on a JASCO FT/IR 4600 Fourier transform infrared spectrophotometer. NMR spectra were recorded on a JEOL ESZ-400S spectrometer, operating at 400 MHz (^{1}H), 100 MHz (^{13}C), and, 396 MHz (^{19}F). Chemical shifts in DMSO-d_6 were reported in the scale relative to DMSO-d_6 (2.50 ppm) for ^{1}H NMR and to $CDCl_3$ (39.52 ppm) for ^{13}C NMR as internal references. Chemical shifts for ^{19}F NMR in DMSO-d_6 was reported in the scale relative to trifluoroacetic acid (-78.50 ppm). ESI mass spectra were measured on a Thermo Scientific Exactive. Dry THF was purified under argon using the Ultimate Solvent System (GlassContour: Nikko Hansen & Co.,Ltd.). A cylinder of CO_2 was purchased from Hokkaido Air Water, Inc.

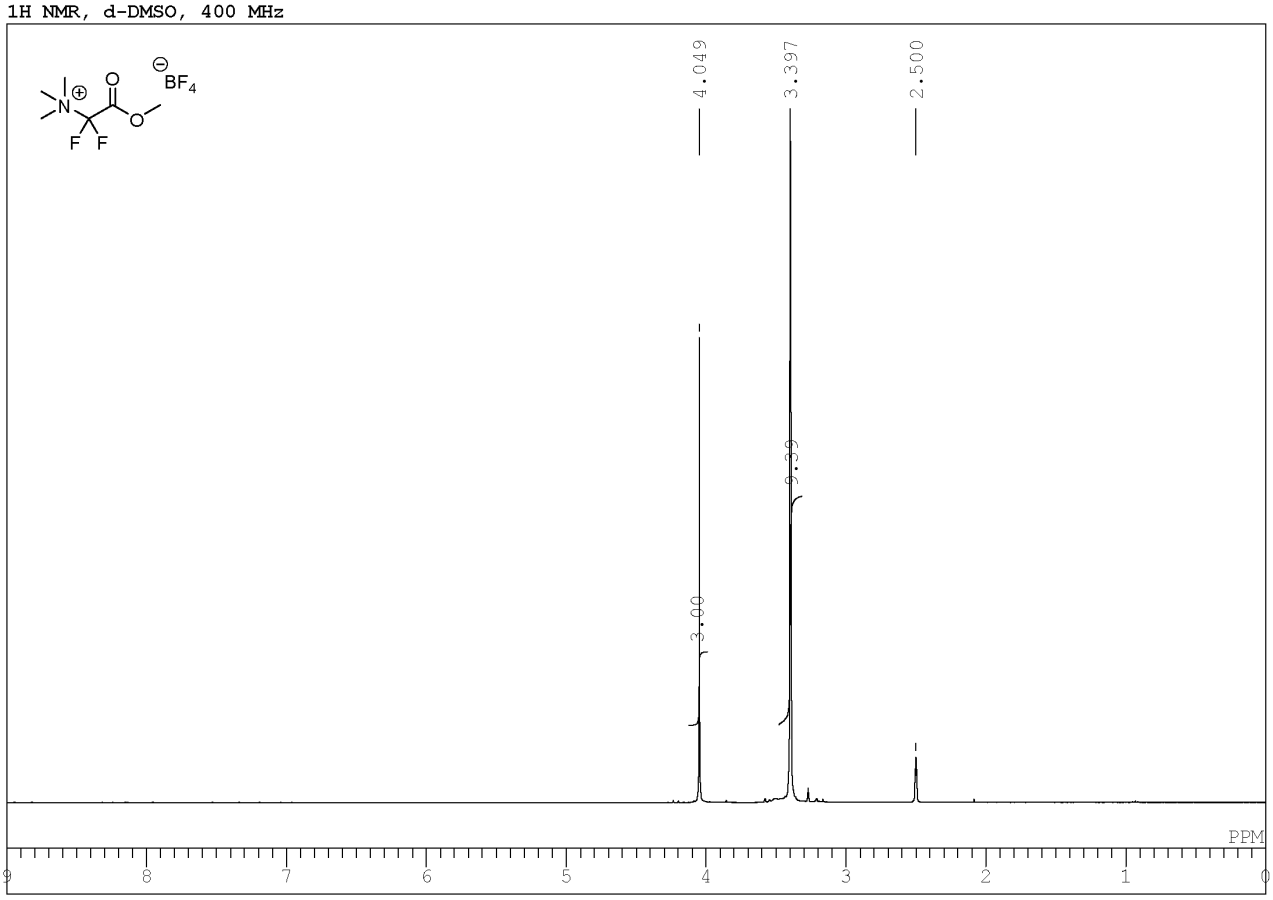
(B) The Synthesis of Difluoroglycine Derivative



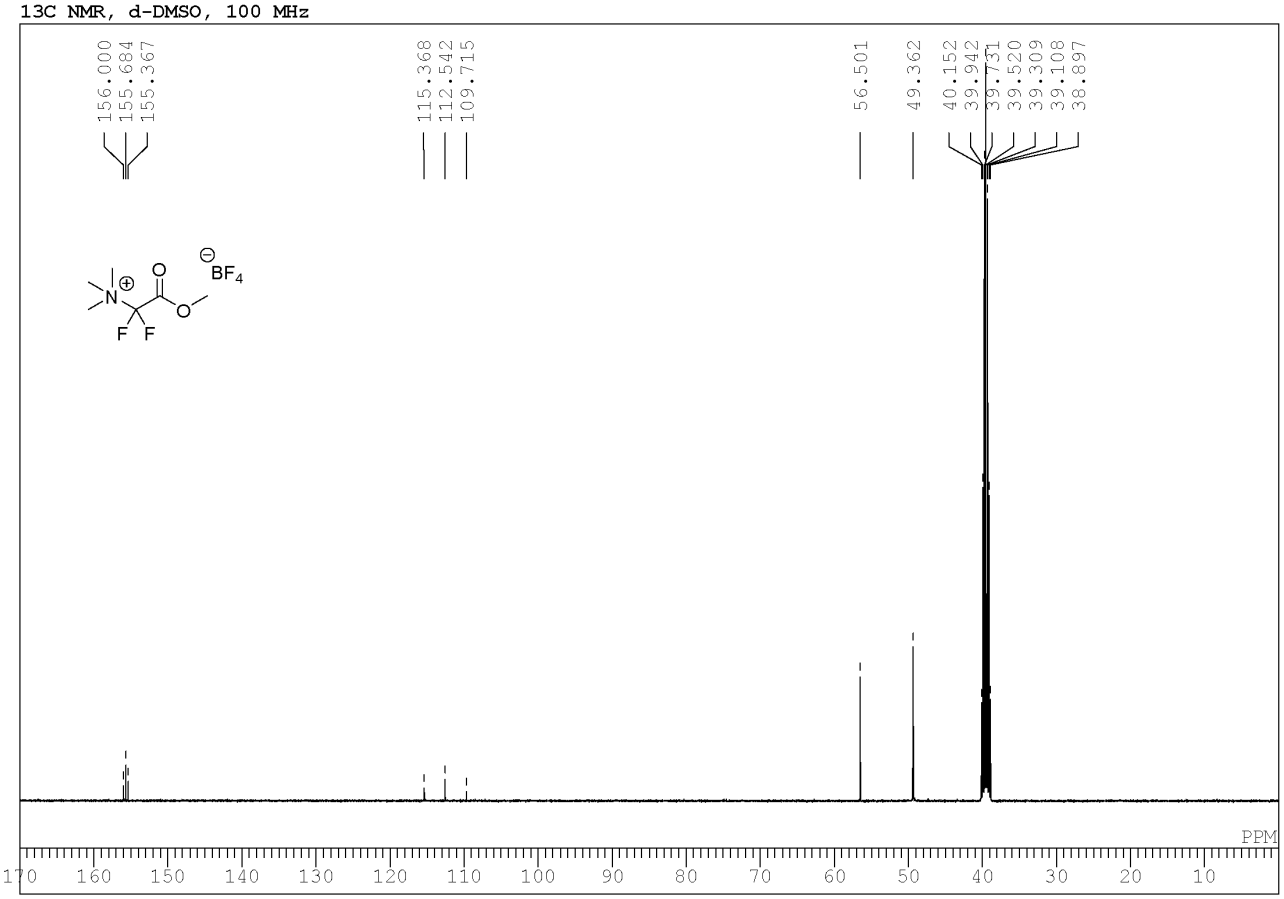
In an oven-dried round-bottom flask was placed TBAT (539.9 mg, 1.0 mmol, 1 equiv). The flask was evacuated and backfilled with CO_2 (3 times) followed by the addition of THF (10 mL). The mixture was stirred at room temperature until TBAT was completely dissolved. After adding trimethylamine (2 M in THF, 500 μL, 1.0 mmol, 1 equiv) at 0 °C, the solution was cooled to −40 °C, and then Me_3SiCF_2Br (156 μL, 1.0 mmol, 1 equiv) was added dropwise. After the resulting slurry was stirred at −40 °C for 20 h, the precipitate was isolated by filtration, washed with hexane, and dried under vacuum. The approximate yield of carboxylate **1** was 96%, as determined by ^{1}H NMR spectroscopy using pyridine as an internal standard. The solids were then treated with the Meerwein reagent (295.8 mg, 2.0 mmol, 2 equiv) in acetone (1.5 mL) for 2 h. After the solvent was removed under vacuum, the resulting solids were washed with a small amount of MeOH to afford methyl carboxylate **3** after filtration (1st crop: 176.6 mg, 2nd and 3rd crops: 28.2 mg, 0.803 mmol, 80% yield).

White solids; IR (ATR): 3068, 2977, 1787, 1483, 1343 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d_6) δ: 4.05 (s, 3H), 3.40 (s, 9H) ppm; ^{13}C NMR (100 MHz, DMSO-d_6) δ: 155.7 (J_{CF} = 31.6 Hz), 112.5 (J_{CF} = 282.7 Hz), 56.5, 49.4 ppm; ^{19}F NMR (376 MHz, DMSO-d_6) δ: −105.4 (CF_2), −151.4 (BF_4) ppm (internal reference: CF_3CO_2H in DMSO-d_6 = −78.5 ppm); HRMS (ESI) *m/z* calcd. for $C_6H_{12}F_2NO_2^+$ [M-BF_4^-]$^+$: 168.0831, found: 168.0833; calcd. for BF_4^- [M-$C_6H_{12}F_2NO_2^+$]: 87.0035, found: 87.0030.

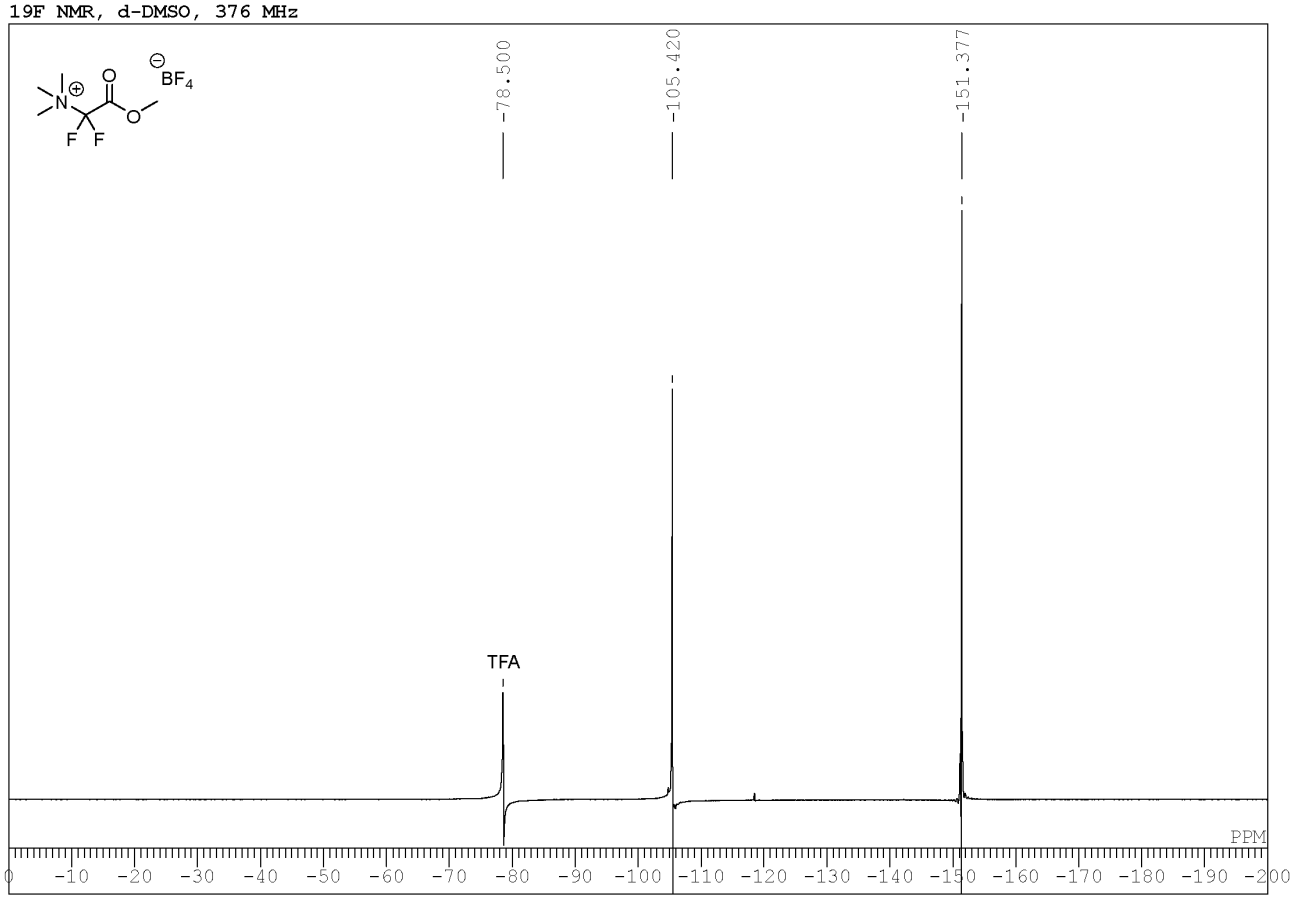
1H NMR, d-DMSO, 400 MHz



13C NMR, d-DMSO, 100 MHz



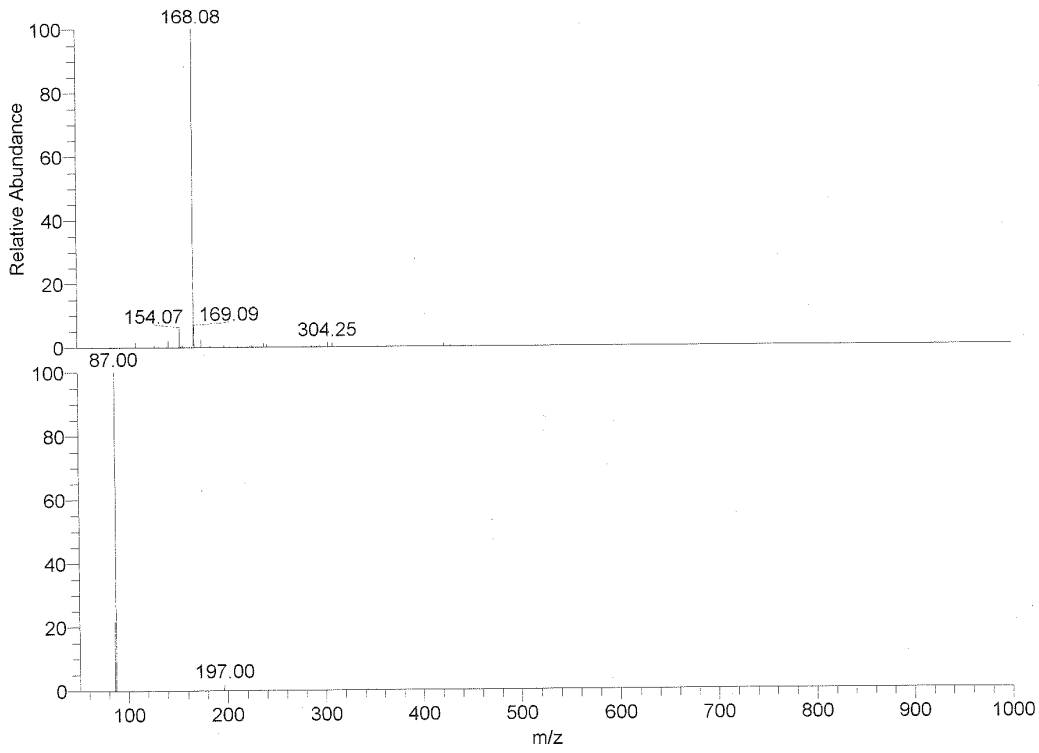
19F NMR, d-DMSO, 376 MHz



Sample No. : C:\Xcalibur\...\1108\BG_191103_Ex_54_pn
Operator name : hayashi harumi
Date : 11/8/2019 10:57:08 AM
Instrumental method : C:\Xcalibur\methods\HESI_100ul\pn_L55_S15_50.meth
Instrumental Analysis Division, Global Facility Center, Creative Research Institution, Hokkaido University

Instrument : Exactive

Mobile phase solvent : MeOH
Sample solvent : MeOH



NL: 6.87E7
BG_191103_Ex_54_pn#
17-25 RT: 0.22-0.31
AV: 5 T: FTMS {1,1} + c
ESI Full ms
[50.00-1000.00]

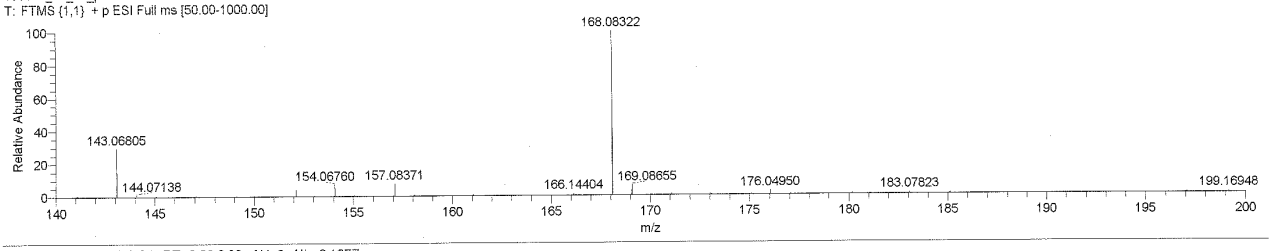
NL: 9.80E7
BG_191103_Ex_54_pn#
18-26 RT: 0.23-0.32
AV: 5 T: FTMS {1,2} - c
ESI Full ms
[50.00-1000.00]

Sample No. : C:\Xcalibur\...\1108\191103_Ex_54_pn
Operator name : hayashi harumi
Date : 11/8/2019 10:12:40 AM
Instrumental method : C:\Xcalibur\methods\HESI_100ul\pn_L55_S15_50.meth
Instrumental Analysis Division, Global Facility Center, Creative Research Institution, Hokkaido University

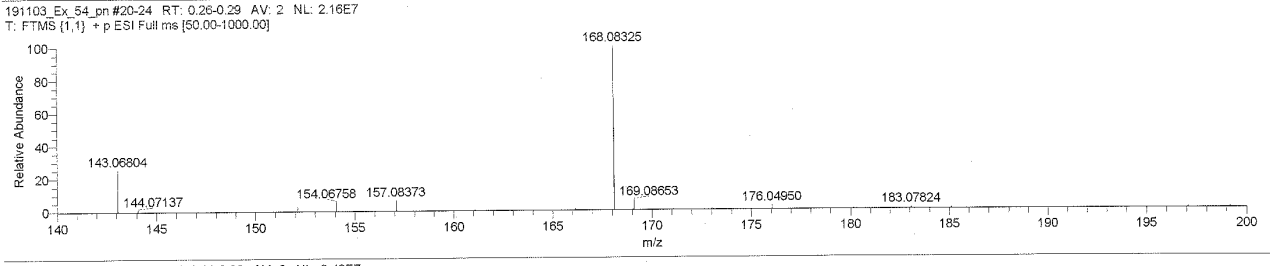
Instrument : Exactive

Mobile phase solvent : MeOH
Sample solvent : MeOH

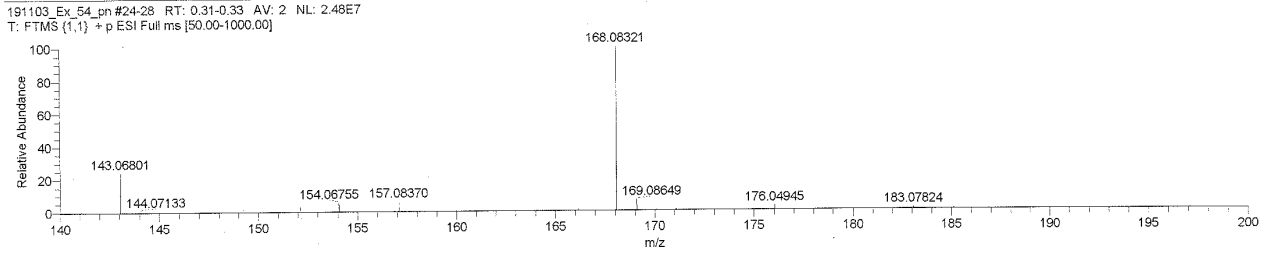
191103_Ex_54_pn #18 RT: 0.24 AV: 1 NL: 1.21E7
T: FTMS {1,1} + p ESI Full ms [50.00-1000.00]



191103_Ex_54_pn #20-24 RT: 0.26-0.29 AV: 2 NL: 2.16E7
T: FTMS {1,1} + p ESI Full ms [50.00-1000.00]



191103_Ex_54_pn #24-28 RT: 0.31-0.33 AV: 2 NL: 2.48E7
T: FTMS {1,1} + p ESI Full ms [50.00-1000.00]

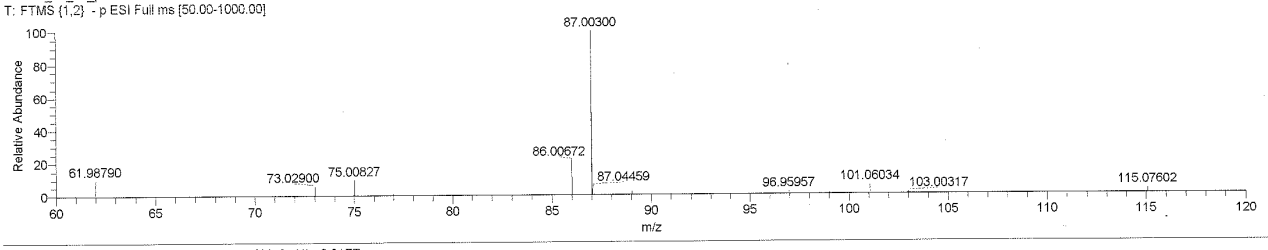


Sample No. : C:\Xcalibur\...\1108\191103_Ex_54_pn
Operator name : hayashi harumi
Date : 11/8/2019 10:12:40 AM
Instrumental method : C:\Xcalibur\methods\HESI_100ul\pn_L55_S15_50.meth
Instrumental Analysis Division, Global Facility Center, Creative Research Institution, Hokkaido University

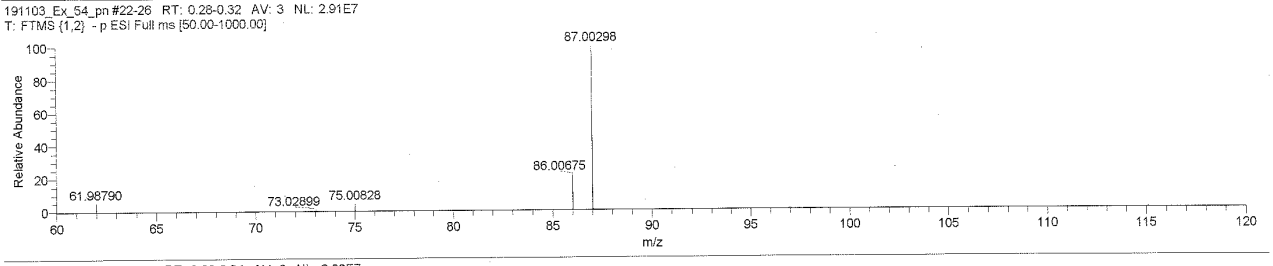
Instrument : Exactive

Neg
Mobile phase solvent : MeOH
Sample solvent : MeOH

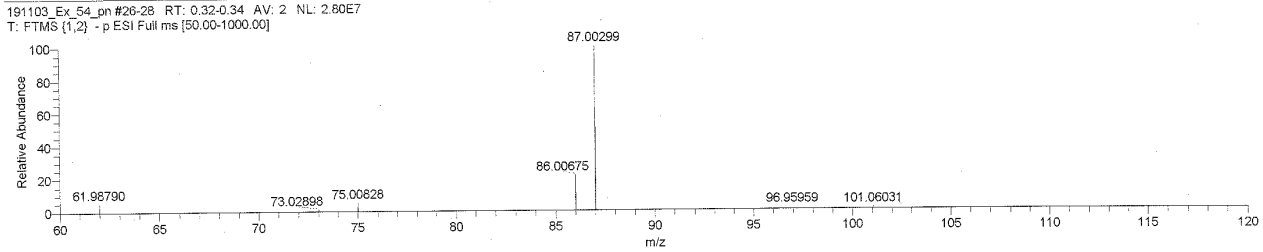
191103_Ex_54_pn #16-19 RT: 0.20-0.23 AV: 2 NL: 8.11E6
T: FTMS {1,2} - p ESI Full ms [50.00-1000.00]



191103_Ex_54_pn #22-26 RT: 0.28-0.32 AV: 3 NL: 2.91E7
T: FTMS {1,2} - p ESI Full ms [50.00-1000.00]



191103_Ex_54_pn #26-28 RT: 0.32-0.34 AV: 2 NL: 2.60E7
T: FTMS {1,2} - p ESI Full ms [50.00-1000.00]



(C) Data for Single Crystal X-ray Structural Analysis

Polymorph	Compound **3**
CCDC Name	CCDC 1971834
Empirical Formula	$C_6H_{12}BF_6NO_2$
Formula Weight	254.98
Crystal System	Orthorhombic
Crystal Size / mm	0.75 × 0.09 × 0.08
a / Å	7.2137(2)
b / Å	11.1355(3)
c / Å	13.0859(3)
α / °	90
β / °	90
γ / °	90
V / Å^3	1051.17(5)
Space Group	$P2_12_12_1$
Z value	4
D_{calc} / g cm^{-3}	1.621
Temperature / K	293
$2\theta_{max}$ / °	10.432 to 152.24
μ (CuK_α) / mm^{-1}	1.643
No. of Reflections Measured	1899
No. of Observations (All reflections)	3624
Residuals: R_1 (I > 2.00σ(I))	0.0610
Residuals: wR_2 (All reflections)	0.1651
Goodness of Fit Indicator (*GOF*)	1.064
Maximum peak in Final Diff. Map / Å^3	0.507
Minimum peak in Final Diff. Map / Å^3	-0.412

(D) Computational Details

To obtain the 30 reactant candidates in Figure 1, we performed an automated search by the single component artificial force induced reaction (SC-AFIR)[1] method starting from α,α-difluoroglycine. SC-AFIR is a method to search for many structures and reaction paths for their interconversions by inducing structural deformations systematically by applying artificial forces between fragments automatically defined in the system. In the search, the model collision energy parameter γ, which defines the strength of the artificial force, was set as 1000 kJ/mol, the maximum number of fragment pairs to

which the artificial force is applied was set as 2 (SC-AFIR2), and the NoBondRearrange option with which the search from local minima having a different bond connectivity pattern from the initial structure is omitted was used to limit the search area.[2] All obtained AFIR paths were reoptimized by the LUP[3,4] method until the highest energy point converged to the first-order saddle point or the number of iterations exceeded 300. As the result, 288 equilibrium (EQ) structures, 309 transition state (TS) structures, and 380 LUP path top (PT) structures, were obtained, where PT is the highest energy structure along a LUP path and regarded as an approximate TS structure.

The reaction path network for CF_3^- + NH_3 + CO_2 shown in Figure 4 (a) and (b) were obtained by the SC-AFIR method combined with the kinetics-based navigation approach. In the search, γ was set as 200 kJ/mol, and the maximum number of fragment pairs to which the artificial force is applied was set as 1 (SC-AFIR1). The kinetics-based navigation is an approach to decide an EQ to which the SC-AFIR procedure is applied next, where the decision is made based on the so-called traffic volume evaluated by the rete constant matrix contraction (RCMC) method. The traffic volume is an index representing the total population influx and outflow that occur in each EQ under given conditions such as initial populations, reaction temperature, and reaction time. Therefore, an EQ having a large traffic volume can be regarded as one contributing to the kinetics significantly, and the kinetics-based navigation preferentially choose such EQs. Further details on the kinetics-based navigation is available in Ref. 5. The search was initiated from 100 initial structures generated by randomly giving mutual positions and orientations between CF_3^- and a NH_3---CO_2 complex, where the most stable NH_3---CO_2 complex prepared separately was used assuming that the reagents for in situ generation of CF_2 was added after mixing NH_3 and CO_2. In the kinetics-based navigation, the initial population 1/100 was given to the 100 initial structures. Three reaction temperatures, 200, 250, and 300 K, were considered, and the highest traffic volume among those obtained with these three temperatures was regarded as the traffic volume of each EQ. The reaction time was set as 1 hour. The search was terminated when a list of EQs with the 3*N* largest traffic volume values was not updated in the last 10*N* path calculations, where *N* was the number of atoms and 11 in this system. All obtained AFIR paths were reoptimized by the LUP method until the number of iterations exceeded 30. Then, LUP paths for transitions taking more than 10^{-10} second were further optimized by the LUP method until the highest energy point converged to the first-order saddle point or the number of iterations exceeded 60. It is known that such an approximate treatment of fast processes affects little on the overall kinetics taking the longer timescale.[6] As the result, 460 EQ structures, 119 TS structures, and 1020 PT structures, were obtained. Population of each EQ after the reaction time of 1 hour was regarded as the yield of each EQ, and the product yield discussed in the main text was obtained by summing yields of all EQs having the corresponding bond connectivity pattern.

The reaction path network for CF_2Br^- + NH_3 + CO_2 shown in Figure 4 (c) and (d) were obtained by the same calculation as had done for CF_3^- + NH_3 + CO_2. As the result, 425 EQ structures, 178 TS structures, and 795 PT structures, were obtained.

The reaction path network for CF_2Br^- + NMe_3 + CO_2 shown in Figure 4 (e) and (f) were also obtained by the same calculation as had done for CF_3^- + NH_3 + CO_2, where $N = 20$ in this case. As the result, 839 EQ structures, 188 TS structures, and 2654 PT structures, were obtained.

1. Maeda, S., Harabuchi, Y., Takagi, M., Saita, K., Suzuki, K., Ichino, T., Sumiya, Y., Sugiyama, K. &

Ono, Y. Implementation and performance of the artificial force induced reaction method in the GRRM17 program. *J. Comput. Chem.* **39**, 233-251 (2018).
2. For these options, see website of the AFIR method: https://afir.sci.hokudai.ac.jp/
3. Choi, C. & Elber, R. Reaction path study of helix formation in tetrapeptides: effect of side chains. *J. Chem. Phys.* **94**, 751-760 (1991).
4. Ayala, P. Y. & Schlegel, H. B. A combined method for determining reaction paths, minima, and transition state geometries. *J. Chem. Phys.* **107**, 375-384 (1997).
5. Sumiya, Y. & Maeda, S. A reaction path network for wohler's urea synthesis. *Chem. Lett.* **48**, 47-50 (2019).
6. Maeda, S., Sugiyama, K., Sumiya, Y., Takagi, M. & Saita, K. Global reaction route mapping for surface adsorbed molecules: a case study for H_2O on Cu(111) surface. *Chem. Lett.* **47**, 396-399 (2018).

(E) Predicted Reactant Candidates

Coordinates of difluoroglycine and obtained 30 reactants with the lowest energy shown in Figure 1. Electronic energies on minima at the wB97XD/6-31+G* Int(Grid=FineGrid) level are indicated in atomic units. The indices of minima, R1, R2, … R30, correspond to those in Figure 1.

difluoroglycine

Energy = -482.793203469575 Hartree

C	-0.400412817137	-0.358246729151	0.768814519728
O	-0.565971386999	0.683907656513	1.574867643848
O	-0.286744406301	-1.504359768219	1.124832827093
C	-0.313379782600	0.055116002438	-0.713700941301
F	1.016478055913	0.413591241203	-0.919707073391
F	-1.013078175109	1.189552787468	-0.956993867132
N	-0.756809206676	-0.974227357103	-1.537426793681
H	-0.565530052945	0.354727093862	2.488966630282
H	-0.414616817502	-1.879732602779	-1.236200027428
H	-0.561278039854	-0.797899206184	-2.516382059488

R1

Energy = -482.814626163586 Hartree

C	0.054775062131	-0.730481675845	2.049854523032
O	-0.374468264226	0.348864468401	2.153452105640
O	0.493324770614	-1.807304812088	1.991696067817
C	-0.356568476754	0.391760378657	-1.198872709965
F	0.990211514201	0.376242456575	-1.002678979730
F	-0.528228915711	0.547438634315	-2.562689091529
N	-0.904250203855	-0.802453340200	-0.700015677631
H	-0.782893952661	1.271769489361	-0.715945019863

H	-1.917112216418	-0.798948651573	-0.749964879150
H	-0.536131946529	-1.614457829578	-1.187765480095

R2

Energy = -482.797075448367 Hartree

C	-0.389298812330	0.091190687955	2.428223389531
O	-0.133290114593	-0.814515535074	1.735452266404
O	-0.631358259388	0.959244847267	3.160179808045
C	-0.100730461892	-1.355349059588	-1.480321168011
F	0.165442315955	-2.279387041424	-2.410417121644
F	-0.962446175435	1.557112380791	0.177942519148
N	-0.483528691829	-0.193719958377	-1.753415838156
H	-0.801326480480	0.951247625138	-0.557188431882
H	0.051744715584	-1.724644727998	-0.471068576937
H	-0.576550664838	-0.008750100618	-2.752315987964

R3

Energy = -482.794810441895 Hartree

C	-0.099328785630	-1.591755468299	1.880463185072
O	-1.060769994497	1.010764647324	0.000234372758
O	0.347509505268	-1.575699815881	2.922269912006
C	-1.381297942188	0.131896812838	-0.767188356172
F	1.524111388048	0.828087643771	-0.093400710416
F	-2.675240837503	-0.166338201329	-0.952824107697
N	-0.588313041126	-0.631214980947	-1.524603326483
H	0.624963197129	1.100399590284	0.096837286874
H	0.413730302051	-0.538101412090	-1.409503776766
H	-0.966706420734	-1.385609697627	-2.075213620657

R4

Energy = -482.797213060850 Hartree

C	-0.137616486102	-1.402367452271	-0.124622638376
O	0.303423435204	0.879461376724	0.705939737763
O	0.190877538026	-2.559630211884	0.025673726107
C	0.525383145977	-0.296270821448	0.731710088231
F	1.423103608121	-0.794455807170	1.556628620070
F	-1.475270226104	1.974834028480	-0.919206794542
N	-1.054442047026	-0.932188017707	-0.988913007284
H	-0.807520169427	1.868669076050	-0.255126538405
H	-1.303019710453	0.048837468936	-1.077863241079
H	-1.526261717455	-1.604460521675	-1.577149093969

R5

Energy = -482.781378055004 Hartree

C	-0.561501393410	-0.067347068459	2.001343652035
O	-1.164932591453	-1.008439656546	1.327350568849
O	-0.585940350595	0.144446491021	3.167630090295
C	0.066670227202	-0.017245780929	-1.912010248594
F	0.313642049238	0.038033013044	-3.224432080595
F	0.173667833782	0.730601100516	1.157158400616
N	-0.653463928073	-0.894206339510	-1.379795171047
H	0.560002886696	0.781350122461	-1.366827699000
H	-0.971486371676	-0.956389073158	0.349097445071
H	-1.038000990925	-1.568373690412	-2.042444099097

R6

Energy = -482.778812795010 Hartree

C	0.501569957861	-0.828706972148	2.653968070911
O	-1.383211854247	-0.575175848335	0.120163514731
O	1.286759227254	-0.838164122086	3.468705631164
C	-0.558017283453	-0.018686222385	-0.789463302188
F	0.635061040323	-0.752274601388	-0.823886850449
F	-0.132799855675	1.223936805914	-0.403890575700
N	-1.191094361357	0.074951978009	-2.032496124781
H	-0.881170401896	-0.724879291237	0.939400565475
H	-1.574941201250	-0.818660913143	-2.316303496659
H	-0.563497896808	0.440088304823	-2.739126573997

R7

Energy = -482.776326585774 Hartree

C	-0.540183932677	-0.597043050686	0.838955660570
O	-1.196003796297	0.415672616183	1.395362951353
O	-0.221297438725	-1.612315689170	1.411177797862
C	0.435584266609	-1.121876741821	-1.133331667116
F	0.715244404812	-0.930521177453	-2.400581137496
F	-1.311573525307	1.954999412899	-1.789638418796
N	-0.288907843501	-0.287675445271	-0.524085525707
H	-1.363450991730	0.187643290373	2.323609019054
H	0.882372795842	-2.027062303451	-0.722941790402
H	-0.973126568244	1.200608206444	-1.321456030780

R8

Energy = -482.765214428561 Hartree

C	0.808980017438	-1.761679259423	1.679575351247
O	-0.478597585198	1.116568112910	0.849941459184
O	1.362309509356	-1.485764454857	2.629451767236
C	-0.302555122050	1.296606115979	-0.302748973293
F	0.828609037604	1.084894246519	-0.948528826837
F	-1.134801745369	1.904698067607	-1.123709539721
N	-1.147121143155	-0.983458719510	-1.143882991366
H	-2.131432773103	-1.072410521716	-0.906454649126
H	-0.631886978217	-1.628553757055	-0.549825876340
H	-1.034845846509	-1.288470712412	-2.106746862492

R9

Energy = -482.773952835216 Hartree

C	-0.851727089433	-0.462508740591	1.124719955388
O	-1.184303962233	0.525433916037	1.945373422534
O	-0.903229604140	-1.640092185930	1.373414611051
C	-0.374048863563	0.032509988358	-0.230077966745
F	0.835916124882	-0.082774918021	-3.684697986057
F	-0.368059968260	1.344999919865	-0.369763572402
N	-0.007805673752	-0.712303023042	-1.168126358914
H	-1.478574359933	0.136661307429	2.786387057823
H	-0.070298813779	-1.695996510475	-0.903702446807
H	0.540789580992	-0.263500635599	-2.796455857339

R10

Energy = -482.760308196684 Hartree

C	0.399429503567	-0.214519900336	2.118443816533
O	0.582257405049	0.366660669275	3.108153621117
O	0.235893771410	-0.840428773398	1.146350108469
C	-0.639188824116	-1.391255623135	-1.896944528381
F	-1.398222273130	1.188141352187	-1.519047574379
F	-0.681612827025	1.992187115930	0.840440750007
N	-0.286908288119	-2.496185458801	-1.987282477225
H	-0.968756828698	1.757799590282	-0.035013780105
H	0.029720120398	-3.445456958323	-2.027200353270
H	-1.133954388542	0.265487104374	-1.670828724235

R11

Energy = -482.765849885398 Hartree

C	0.691261994700	-1.306963296442	-0.235805425021

O	-1.509144166900	2.105491194794	1.483979057983
O	0.394316124756	-0.395186932180	0.653198148064
C	-2.363572777764	2.794582787315	1.193969810115
F	1.777480800617	-1.971149697613	0.129072175333
F	-1.722875066649	0.223264821264	-0.972176454965
N	0.049922438348	-1.544362540295	-1.303861436426
H	-0.413292637551	0.085022622073	0.362120218981
H	0.438081629694	-2.293282284925	-1.866617022392
H	-1.203520968366	-0.514987556006	-1.366808213069

R12

Energy = -482.762355318733 Hartree

C	-0.551295903810	-0.245049708697	2.331024414277
O	-0.741327122415	-1.077557865248	1.529619823546
O	-0.382148765227	0.527125077023	3.178142403620
C	-0.204762771659	-0.021546441771	-1.722197690143
F	-0.063018192460	0.133445281103	-3.040928711974
F	0.099345829973	1.511363524468	0.410043452971
N	-0.581062781376	-1.246295016604	-1.465621302741
H	0.021400499861	1.046361459043	-0.465231448617
H	-0.724558479905	-1.496400676918	-0.492090654709
H	-0.733914942207	-1.949016514357	-2.185689427710

R13

Energy = -482.764317151512 Hartree

C	-1.395414041104	-0.309922280105	1.375814057524
O	-2.117869229846	-0.325084205208	2.494887796226
O	-0.272449785029	-0.747827196699	1.282006193404
C	0.764378251252	-0.110641483798	-2.470570775056
F	1.995099288622	-0.607264260972	-2.377969011082
F	0.582993589562	0.303525425232	-3.705969031398
N	-0.097122955104	-0.029191409806	-1.583596111743
H	-1.584693337315	-0.741676174387	3.193985580778
H	0.216216109713	-0.400973030805	-0.683775000391
H	-1.952480519862	0.151483734534	0.552257160318

R14

Energy = -482.760568478488 Hartree

C	-0.816981391116	-1.160596524877	1.423057298049
O	-0.808010134806	0.127867944959	1.697256413723
O	-1.047551017002	-2.041370579930	2.198773568165

C	-0.501179154897	-1.471449676102	-0.003707225114
F	0.253377680349	0.379864172193	-2.721477120118
F	-0.208027584064	1.599478651864	-0.508797091937
N	-0.251811445992	-1.673309378459	-1.113332869924
H	-0.594355600348	0.691400060973	0.919169351736
H	0.130507814934	-0.523791330117	-2.430369162909
H	-0.017311796291	1.254335777531	-1.383502303125

R15

Energy = -482.762198666521 Hartree

C	0.350766125307	-0.231720807310	1.016233445062
O	-0.732995526138	0.316004195673	1.331621673749
O	1.372106566208	-0.483675804199	1.623165685770
C	0.412900729125	-0.700438353550	-0.498949984281
F	1.507289013152	-1.268974179745	-0.902899474859
F	-2.553857804915	0.601846615464	-0.231235442088
N	-0.520647855375	-0.573751501032	-1.372462214257
H	-1.881966999624	0.551784516203	0.531710606661
H	-1.425517017168	-0.120238256902	-1.097046769795
H	-0.389419859772	-0.908407306561	-2.323066667429

R16

Energy = -482.753221614938 Hartree

C	0.572463353152	-0.711661517260	-0.534380180783
O	-0.734373857390	1.051203403213	0.972638706836
O	1.088032887577	-1.418438115369	0.268365960751
C	-0.833094472349	-0.898106096797	-1.058065762162
F	1.218610784967	0.294994377622	-1.124665897248
F	-1.224500793983	0.074729995583	-1.876465297370
N	-1.589472858339	-1.856556581199	-0.798978272448
H	-1.005235260275	1.905960868493	0.617008855152
H	-1.132873366188	-2.502442142191	-0.154718036930
H	-0.220899046407	1.242744926012	1.766330782913

R17

Energy = -482.736998341643 Hartree

C	0.150379167545	2.533544693338	1.099753803691
O	-1.569674077547	-2.373467214714	0.272759304340
O	-0.665485979621	3.320683381372	1.055391715433
C	0.514277429716	-0.049384280226	-1.290633294425
F	1.222381642640	-1.165256437937	-1.089483466574

F	1.312937106833	0.837716684259	-1.845623271057
N	-0.675040725964	0.158019537744	-1.018301639842
H	-2.218100924021	-2.483638294107	0.977626936865
H	-0.819607008171	-2.933653881356	0.505014400593
H	-1.113409260634	-0.662135070402	-0.589433630499

R18

Energy = -482.748315896604 Hartree

C	-0.217277764187	-1.913727146999	0.819647475506
O	-0.013641585634	-1.065231846257	1.760174238379
O	-0.248899031223	1.201364645610	0.336141227727
C	-0.488213064179	1.076411391976	-0.846151952591
F	-0.133442893058	-3.119354396966	1.368023462537
F	-0.620592313953	2.167402893742	-1.618106890688
N	-0.650083039152	-0.048747512112	-1.539477453088
H	-0.074803658190	-0.155985766825	1.354870064980
H	-0.566062401116	-0.935984552884	-1.030536930974
H	-0.848326878510	-0.023718591227	-2.527512383266

R19

Energy = -482.737619100570 Hartree

C	-0.601052960466	0.091671279171	1.121788797195
O	-1.716230549422	-0.646528521721	1.078785091981
O	0.102925637823	0.218958314390	2.090947671802
C	0.539771797652	0.504150866590	-0.837538751711
F	1.437805114130	-0.446811522183	-0.693082589776
F	0.786806001508	1.172466374703	-1.932363302533
N	-0.421634647358	0.760579968357	-0.095225023681
H	-1.826152353691	-1.048240296563	1.955118591010
H	-0.814660358274	-1.968828286027	-2.324512986746
H	-1.348920311129	-1.454989058662	-2.286846639016

R20

Energy = -482.741700046415 Hartree

C	0.652082911649	-0.055649567707	0.609476787050
O	-2.449652283058	-0.223567168914	0.929676327933
O	1.826578254132	-0.158161327432	0.845186483213
C	0.379474404293	-0.352585349838	-1.689894259755
F	-0.142755539142	-0.927249809139	-2.744220243942
F	1.260826632474	0.541271639640	-2.074861150640
N	0.028205635680	-0.626890343549	-0.528560030573

H	-3.282828122157	-0.631868194862	1.189540694180
H	-0.078340474278	0.419048347049	1.278738460225
H	-2.054934048803	-0.801919107239	0.261987790860

R21

Energy = -482.744866113294 Hartree

C	-0.236101539365	0.311548010933	1.866685134741
O	-1.003477089662	0.959458237962	1.048923509189
O	-0.021612618484	0.487778670203	3.019731612721
C	-0.872449239855	-0.164608343514	-1.464368704179
F	-1.481900675882	0.153279892166	-2.608144024091
F	0.398995122825	-0.740915001259	1.207012535441
N	-0.093594844801	-1.196219667956	-1.660510293002
H	-0.991895288471	0.576330800073	0.092987401307
H	0.430771099147	-1.542391986979	-0.865529509132
H	0.009922445331	-1.661831493598	-2.559716804482

R22

Energy = -482.734701613681 Hartree

C	-0.084757338449	0.298254718212	1.484551899565
O	-1.279558976311	0.852046529661	1.223675118835
O	0.593488865186	0.031501104880	2.419084398502
C	-0.451174933706	0.376321632578	0.080031701728
F	0.887074396295	-1.891573024271	-2.394368594680
F	-0.020214864676	1.393526086421	-0.671817550821
N	-0.919470853974	-0.669295931641	-0.723081130288
H	-1.273335765182	-1.448631810581	-0.175857006947
H	0.324449888899	-1.412706829348	-1.795013323604
H	-1.637843047291	-0.347013357861	-1.370134653768

R23

Energy = -482.732778807392 Hartree

C	-1.115116946685	0.140767370471	0.866418795530
O	-1.510913748947	-0.877257369469	1.641855650147
O	-1.349699293917	1.299736217515	1.046836937013
C	0.248641567534	0.184268313896	-1.238705226295
F	1.375860262247	-1.773295005007	-2.694397543683
F	0.106279467736	1.486316966908	-1.198449245650
N	-0.365919018725	-0.407514717207	-0.210397709713
H	-2.023414880501	-0.509173123283	2.379024739031
H	1.046463892088	-0.942266076391	-2.307144212873

H	-0.273523929950	-1.419153459477	-0.207971324922

R24

Energy = -482.724978283980 Hartree

C	-1.117680748457	-0.406142208190	-0.400452239590
O	-0.796313226307	-1.215106189235	0.618531540749
O	-0.339132226429	0.397932605455	-0.898268211687
C	1.809883917121	0.081480641815	0.458639724425
F	2.599080668818	-0.111515201204	-0.558993234842
F	1.963459830752	1.333494434992	0.781806127518
N	-2.395920822321	-0.581156738913	-0.806590551773
H	0.139538771655	-1.008432044599	0.838424402255
H	-2.736065686790	-0.036590181013	-1.581710899330
H	-2.988193107172	-1.271536001131	-0.374315799187

R25

Energy = -482.724300210028 Hartree

C	-0.174337224603	-0.035529683576	2.118533289647
O	-1.248907083059	-0.365576500570	1.783283409340
O	0.850129393199	0.288148743556	2.556429991118
C	0.376397146221	0.001613051707	-0.731875370597
F	1.153500397628	-0.415005138390	-1.809587529955
F	-0.022496920493	1.283130868446	-1.124471627781
N	-0.984155066810	-0.766321430530	-1.130642476535
H	-1.685557339598	-0.526904213640	-0.428277570076
H	-0.811228869758	-1.770102025898	-1.088138930831
H	-1.314687061938	-0.511024553058	-2.068182325804

R26

Energy = -482.710546064105 Hartree

C	-0.114935572271	-0.121718425870	2.329813452850
O	-0.974595785863	-0.768002744763	1.879655154182
O	0.735726135531	0.516910365356	2.801628680052
C	0.742579966321	0.661126759105	-0.621626508054
F	1.621370034428	0.078270256649	-1.392485110702
F	0.277658396573	1.646428626729	-1.342488668541
N	-1.732900447439	-1.366776552233	-1.431642637877
H	-1.588906395881	-1.229836352719	-0.436417558856
H	-0.892780140063	-1.760020419173	-1.841296391845
H	-1.934558820549	-0.473952395032	-1.868069552678

R27

Energy = -482.718950903638 Hartree

C	0.267629918934	-0.151693042990	1.184237936626
O	-0.388915800334	-0.803753594717	1.922108839212
O	1.368342890770	0.647101883896	1.233840120413
C	0.716566700118	0.457168610064	-0.041658526583
F	1.452879908236	-0.164275999885	-0.944823542013
F	0.192479770139	1.555175536406	-0.554925828054
N	-1.635063538074	-0.948784596456	-1.107563325148
H	-2.473709486299	-0.515357212018	-0.730495946252
H	-1.619519971184	-1.906229242707	-0.766235496357
H	-1.742033021530	-0.986923223546	-2.117413373289

R28

Energy = -482.714239391937 Hartree

C	0.109791730373	1.277575712858	0.248129648913
O	-1.672019604950	-2.330960159952	0.227810072532
O	0.073657353725	2.012712501364	1.193004889552
C	0.655348437243	0.741102972679	-1.024130730209
F	1.821062440564	0.061702165495	-1.065365852924
F	0.434509267421	1.234997582583	-2.256260527555
N	-0.432163367726	0.179342574953	-0.322014257553
H	-2.634014450401	-2.391225870045	0.268414438153
H	-0.882744008097	-0.730508232603	-0.144267844794
H	-1.334770427373	-2.872310129285	0.951751022382

R29

Energy = -482.700362053186 Hartree

C	-0.579202735555	-0.530532079091	1.719258652443
O	-0.437918887803	-0.676716775350	3.032122147864
O	0.619906619231	-0.292221872710	1.252387746733
C	-0.343270623587	0.004234403062	-2.579267958675
F	0.780492433139	0.258810606127	-3.198375310343
F	-1.320812331217	-0.060852582707	-3.465618609564
N	-0.425810839646	-0.144432076340	-1.347684215645
H	-1.315436274259	-0.853807871616	3.392202227268
H	-1.366375219418	-0.337616769975	-1.009264578696
H	0.527085229922	-0.184435863359	0.281310757162

R30

Energy = -482.676595442259 Hartree

C	-0.801055085218	-1.065123076873	0.287002021833
O	-0.877370152070	0.218444931465	0.658552481230
O	0.498097397500	-1.409837679507	0.269557218240
C	0.553712361881	1.505523620391	-1.498153778290
F	1.642773831192	1.675024660379	-0.798946908107
F	-0.160901396997	2.576174175222	-1.246911409173
N	-1.841340446324	-1.742238651107	0.023083826456
H	-1.818675228722	0.446953553381	0.664761392927
H	0.591530079413	-2.319134204009	-0.039331692046
H	-1.648113989851	-2.703358211323	-0.242542294546