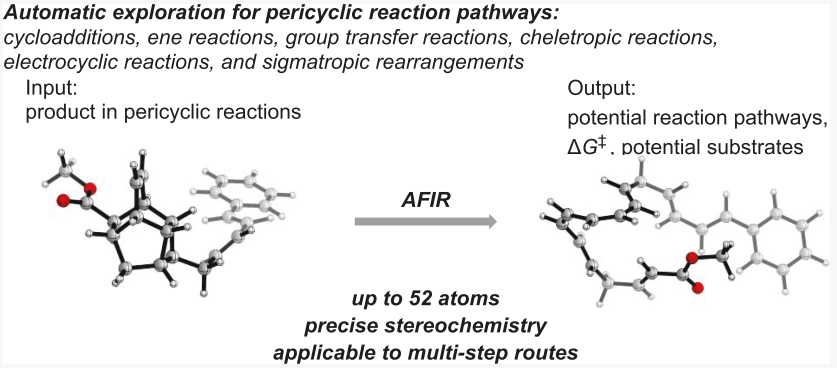
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Prediction of High-Yielding Single-Step or Cascade Pericyclic Reactions for the Synthesis of Complex Synthetic Targets

Tsuyoshi Mita,* Hideaki Takano, Hiroki Hayashi, Wataru Kanna, Yu Harabuchi, K. N. Houk, and Satoshi Maeda*

Cite This: *J. Am. Chem. Soc.* 2022, 144, 22985−23000
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

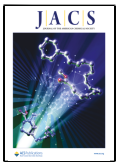


ABSTRACT: Pericyclic reactions, which involve cyclic concerted transition states without ionic or radical intermediates, have been extensively studied since their definition in the 1960s, and the famous Woodward−Hoffmann rules predict their stereoselectivity and chemoselectivity. Here, we describe the application of a fully automated reaction-path search method, that is, the artificial force induced reaction (AFIR), to trace an input compound back to reasonable starting materials through thermally allowed pericyclic reactions via product-based quantum-chemistry-aided retrosynthetic analysis (QCaRA) without using any a priori experimental knowledge. All categories of pericyclic reactions, including cycloadditions, ene reactions, group-transfer, cheletropic, electrocyclic, and sigmatropic reactions, were successfully traced back via concerted reaction pathways, and starting materials were computationally obtained with the correct stereochemistry. Furthermore, AFIR was used to predict whether the identified reaction pathway can be expected to occur in good yield relative to other possible reactions of the identified starting material. In order to showcase its practical utility, this state-of-the-art technology was also applied to the retrosynthetic analysis of a natural product with a relatively high number of atoms (52 atoms: endiandric acid C methyl ester), which was first synthesized by Nicolaou in 1982 and provided the corresponding starting polyenes with the correct stereospecificity via three pericyclic reaction cascades (one Diels−Alder reaction as well as 6π and 8π electrocyclic reactions). Moreover, not only systems that obey the Woodward−Hoffmann rules but also systems that violate these rules, such as those recently calculated by Houk, can be retrosynthesized accurately.

INTRODUCTION

Pericyclic reactions, wherein bond-forming and -cleaving events occur simultaneously via cyclic transition states, have emerged as the third major class of reactions in organic chemistry,[1] in addition to conventional ionic and radical reactions. These concerted reaction systems have been intensively studied by Woodward and Hoffmann, who asserted the conservation of orbital symmetry between the reactants and products in the so-called Woodward−Hoffmann rules.[2] Dewar and Zimmerman interpreted these rules in a more comprehensive way, that is, those pericyclic reactions via an aromatic transition state are thermally allowed, while reactions via an anti-aromatic transition state are photochemically allowed (Dewar−Zimmerman model).[3] There are several types of pericyclic reactions, including cycloadditions, ene reactions, group-transfers, cheletropic, electrocyclic, and sigmatropic reactions, which are all kinetically governed by the Woodward−Hoffmann rules. The activation barrier for the forbidden process is significantly higher than that for the allowed pathway; both these values can now be precisely calculated computationally,[4] confirming the robustness of the Woodward−Hoffmann rules discovered more than 50 years ago.[5] In particular, Houk and co-workers significantly

Received: September 14, 2022
Published: December 1, 2022



https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

contributed to this field by using quantum-chemical calculations to establish a general theory of substituent effects on the reaction rates and stereoselectivity of pericyclic reactions.[6] In addition, the transition-state structures for all categories of pericyclic reactions were systematically calculated using ab initio quantum-chemistry methods.[4]

Given the recent significant improvement of the CPU capabilities of computers, quantum-chemical calculations have become very popular for unveiling the reaction mechanisms of reported transformations as well as newly developed reactions. Not only computational but also organic chemists now use DFT calculations frequently to understand how reactions work. However, DFT calculations usually require a rough idea of the overall reaction processes, and appropriate three-dimensional structures should be considered beforehand. Thus, the development of a fully automated reaction-path search that can precisely determine equilibrium structures (EQs) and transition states (TSs) without relying on input structures would be extremely desirable. To this end, various automated reaction-path search methods,[7] such as the gradient extremal and eigenvector following methods,[7a] the anharmonic downward distortion following method,[7b] the single/double ended growing string method with systematic molecular graph transformations,[7d,h] the freezing string method with the Berny algorithm,[7g] the nanoreactor,[7l] and the KinBot,[7m] have been developed to date.[8] Despite their successful performance in benchmark cases, only a few of these methods have already found practical applications. In part, this is because care is needed when a method relying on empirical templates and/or rules is applied to a new system. It is also partially due to the lack of exhaustiveness of the individual methods.[9]

The artificial force induced reaction (AFIR) method,[7c,8a,10] a DFT-based automated reaction-path search method, can automatically output reaction pathways by using appropriate search options, such as the collision energy parameter (γ) and the DFT calculation level, without using any empirical templates or rules. One strong advantage of AFIR is that it can search for transformations of any type by a single algorithm, that is, applying a force between fragments. In other words, AFIR finds pathways of various types without combining any other algorithms such as conformational search algorithms. Moreover, the AFIR method has been proven to be able to find even high-energy compounds exhaustively,[10,11] enabling the use of retrosynthetic disconnection to search for unstable compounds by pushing and pulling an artificial force between atoms along the potential energy surface in a suitable manner. A series of AFIR searches with DFT calculations can be automatically carried out as a single operation under quantum-chemistry-aided retrosynthetic analysis (QCaRA) conditions using a developmental version of the Global Reaction Route Mapping (GRRM) program.

QCaRA is a reaction-discovery concept that has been proposed in 2013 to explore potential reactants starting from a target product using an automated reaction-path search method.[8a] However, its application had been limited until very recently to one-step reactions. Even with such a severe restriction, it had been successfully used for the discovery of a real chemical reaction in 2020 (vide infra).[12a,b] This severe limitation was due to the exponential combinatorial increase of the search space for potential reactants. Recently, we have found a way to circumvent this limitation by employing an inverse kinetic simulation algorithm during an automated reaction-path exploration.[13] In principle, any automated reaction-path search method that is able to exhaustively explore reactant species that are much higher in energy than the target product can be used in QCaRA. The additional inverse kinetic simulation is required to narrow down the search space according to kinetics-based screening. Hitherto, QCaRA for multistep reactions has been achieved only by AFIR in combination with the inverse kinetic simulation.[13]

Because of the very high number of structures that must be searched in QCaRA, the number of atoms in systems to which QCaRA has been applied has so far been limited. Nevertheless, even such small systems can still be expected to be of high potential utility in organic chemistry to design new chemical reactions. Indeed, using QCaRA, we successfully searched for reactants of several useful organic molecules. Difluoroglycine (10 atoms),[12] α-amino nitrile (19 atoms),[13] and α-formyloxyamide (20 atoms)[13] have already been computationally retrosynthesized into reasonable reactants via the carboxylation of ammonium ylide, the Strecker reaction of acetone, and the Passerini reaction of formaldehyde, respectively. In particular, one of the reactant combinations proposed by QCaRA for difluoroglycine, which consisted of a *tert*-amine, difluorocarbene, and CO_2, was tested experimentally to afford a difluoroglycine derivative in high yields.[12a] Although these demonstrations are highly accurate and useful for the development of new chemical reactions, they remain limited to relatively simple, nonstereoselective reactions.

In this paper, we report a stereospecific QCaRA for a wide variety of known pericyclic reactions that provide simple molecules and even complicated ones, such as, for example, endiandric acid C methyl ester (Figure 1a). There are several

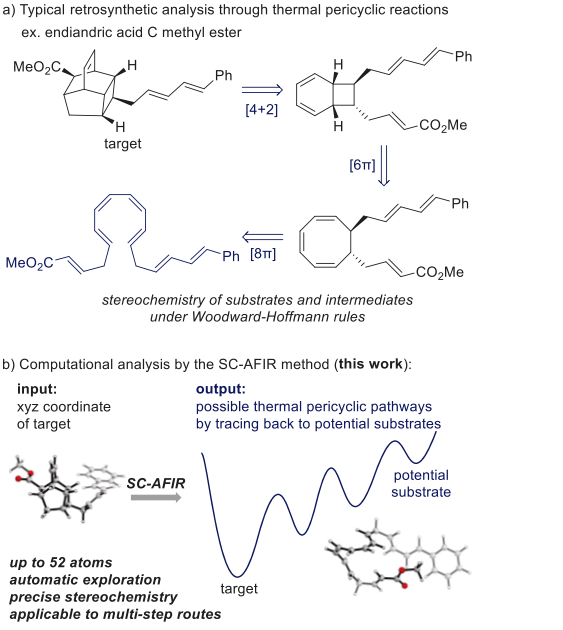


Figure 1. Retrosynthetic analysis involving thermal pericyclic reactions.

computational advantages for the calculations of pericyclic reactions over common ionic and radical pathways: (1) Pericyclic reactions are perfectly atom-economical processes, and thus can be carried out under heating conditions without additives, which we assume will allow for a correct prediction of the stereochemistry; (2) solvent effects should be minimal in concerted pericyclic reactions; thus, predictions for real

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

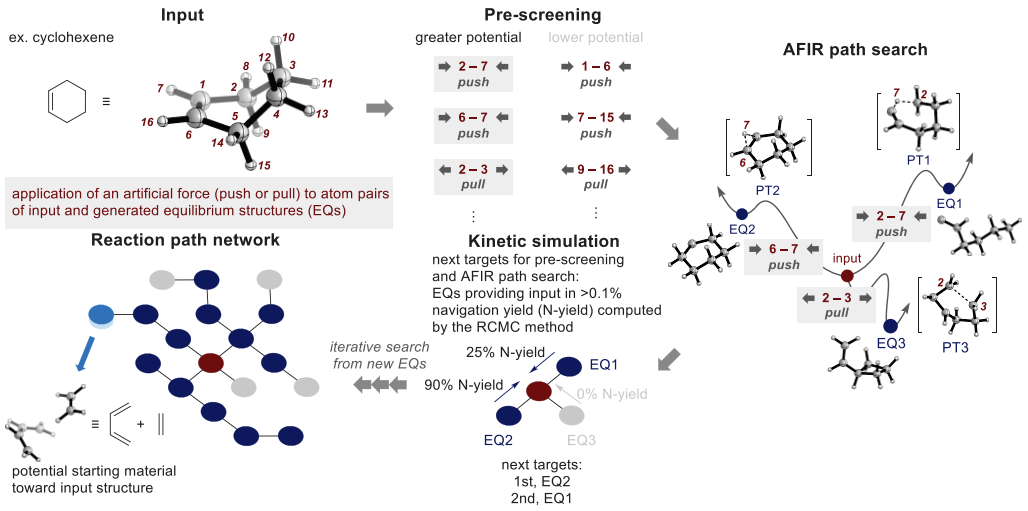


Figure 2. Systematic process for the generation of the reaction path network by the AFIR method.

experimental systems could be made without taking the solvent effect into account in the calculations (gas phase conditions),[4] which would greatly reduce the calculation cost; (3) the stereospecificity for the forward and backward cyclization should be the same (conrotatory/disrotatory additions and suprafacial/antarafacial addition), allowing the stereoselectivity of the forward pathways to be confirmed during the backward path search. Considering these calculation advantages, we anticipated that AFIR searches with DFT calculations would utilize the appropriate stereospecificity, affording reasonable starting compounds with the correct stereochemistry without the need for any previous experimental data (Figure 1b). Here, we report the results of fully automated AFIR calculations of all the reported types of pericyclic reactions, that is, cycloaddition, ene, group-transfer, cheletropic, and electrocyclic reactions as well as sigmatropic rearrangements, under QCaRA conditions. Moreover, the QCaRA of a 52-atom natural product, endiandric acid C methyl ester, which was synthesized in a single step by Nicolaou in 1982[14] based on the assumed biosynthetic route,[15] successfully traces the reaction back to the corresponding starting polyene via three continuous retro-pericyclic reactions without using human expertise or intuition with respect to this specific pericyclic reaction cascade. Moreover, in addition to systems that obey the Woodward–Hoffmann rules, systems that violate these rules, such as the conrotatory 14π electrocyclic reaction system recently calculated by Houk,[16] were investigated.

MATERIALS AND METHODS

Automated Reaction-Path Search Method. We conducted all QCaRA calculations using the GRRM program[8a] combined with the Gaussian 16 program[17] with the unified options (for details on the calculation parameters, see the Supporting Information). All calculations were conducted using the single component artificial force induced reaction (SC-AFIR) method with the ωB97X-D functional in combination with the Def2-SV(P) (for S) or the SV (for all other atoms) basis sets with Grid = FineGrid (Figure 2).[18] Gibbs energy correction evaluated under the harmonic vibrational and rigid-rotor approximations was applied by setting all frequencies (except for imaginary ones) that are less than 50 cm^{-1} to 50 cm^{-1}.[19] The AFIR calculation is stopped after 1000 reaction pathways are found. To efficiently identify the potential uphill local minima from the input structure (cyclohexene is shown as an example in Figure 2), the calculations were performed by following the process of prescreening, AFIR path search, and kinetic simulation to construct the reaction path network. Prescreening calculations were conducted with regard to atom pairs by applying the artificial force (both push and pull forces), which allowed us to search for chemically important reaction pathways such as bond-forming or -cleaving pathways, preferentially (pathways with greater potential) rather than those of conformational change.[10] Then, the automated AFIR path search was performed with the atom pairs according to the prescreening calculations to explore the possible pathways, including the corresponding path tops (PTs; *vide infra*) and EQs. After the path-search process, kinetic simulations were conducted by using the rate constant matrix contraction (RCMC) method,[20] which dictates the direction of the path search based on the kinetic accessibility from each EQ to the input structure in general regarding the input structure as the reference structure. However, when a structure that is more likely a product than the input structure emerges during the QCaRA calculation, it is automatically set as the reference structure. Such a case happens when the search obtains a kinetically feasible path from the input structure to a more stable structure. In this kinetic simulation, we calculated the navigation yield (N-yield) of each EQ, defined as the yield of the above reference structure (not always the input structure) in the obtained reaction map, which can be accessed from the corresponding EQ via the generated reaction pathway at each temperature (300, 400, or 500 K). EQs that are expected to provide the reference structure in less than 0.1% N-yield in the RCMC method were not considered for further calculations. We envisioned that by using this systematic and iterative search method, a completely nonempirical process could be developed to obtain a set of pericyclic reaction pathways by tracing back from the energetically stable structures. A reaction path network is eventually created in which EQs with finite N-yields should be possible reactants.

In this search process, the obtained AFIR paths tend to pass through near-TSs of the corresponding reaction pathways (AFIR paths), giving approximate TSs (PTs). Such PTs can be readily re-optimized to find the corresponding real TSs by the locally updated planes (LUP) calculations (LUP paths), followed by SADDLE calculations.[21] In this study, the LUP and SADDLE calculations of the lowest PT among several conformational isomers were carried out at the ωB97X-D/6-31G(d,p) level with Grid = UltraFine to afford first-order saddle-point structures. All TSs were also confirmed using intrinsic reaction coordinate (IRC) calculations,[22] and connected reactants and products were obtained. Finally, single-point calculations were performed at the ωB97X-D/6-311+G(d,p) level with Grid = UltraFine. The diagram thus obtained is shown in each figure. All calculations were performed with a charge of zero and a singlet spin multiplicity in closed-shell, single Slater determinant spin-restricted method, and thus did not consider the diradical pathways that have often been suggested by Houk, especially for thermally

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

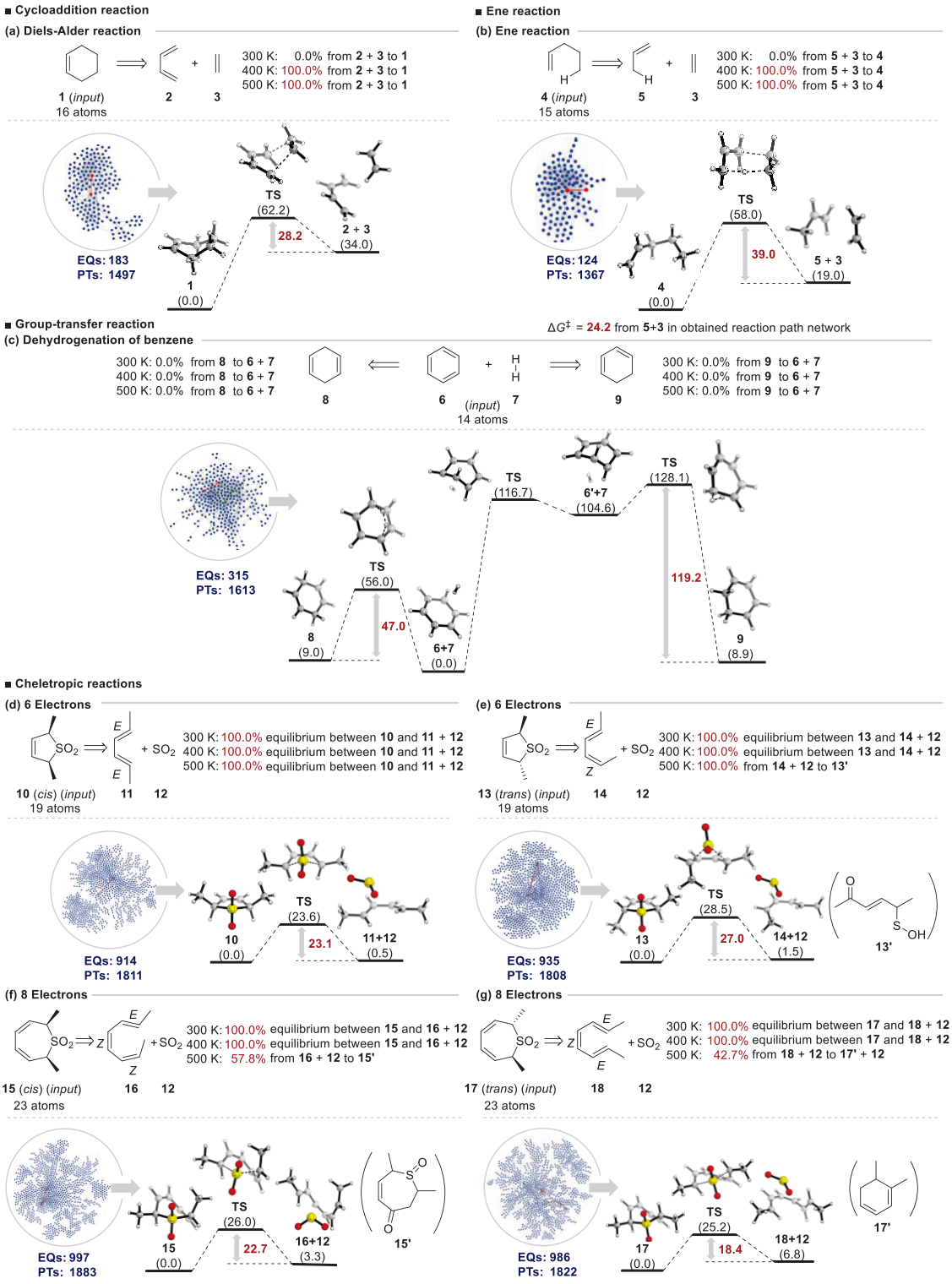


Figure 3. QCaRA for pericyclic reactions-1 (Diels-Alder, ene, group-transfer, and cheletropic reactions). AFIR search: ωB97X-D/Def2-SV(P) for S atoms and SV for other atoms; structure refinement: ωB97X-D/6-311+G(d,p)//ωB97X-D/6-31G(d,p).

forbidden pericyclic reactions.[4,6] The obtained energy diagrams did not fully consider conformations of obtained structures on the networks because only one PT was re-optimized by LUP and SADDLE calculations; however, the re-optimization of all PTs should be prohibitively time-consuming and energy diagrams based on limited LUP optimization can be expected to allow a sufficient rationalization of the results of the automated reaction path search. The red lines in the obtained reaction path networks represent the desired pathways from the product to the reactant.

■ RESULTS AND DISCUSSION

Diels-Alder, Ene, and Group-Transfer Reactions. Prior to the studies of stereospecific pericyclic reactions, we first investigated retrosynthetic pathways for the simple Diels–Alder product 1-cyclohexene (**1**) (Figure 3a; 16 atoms). QCaRA automatically generated 183 EQs and 1497 PTs to produce a reaction path network indicating the associative

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

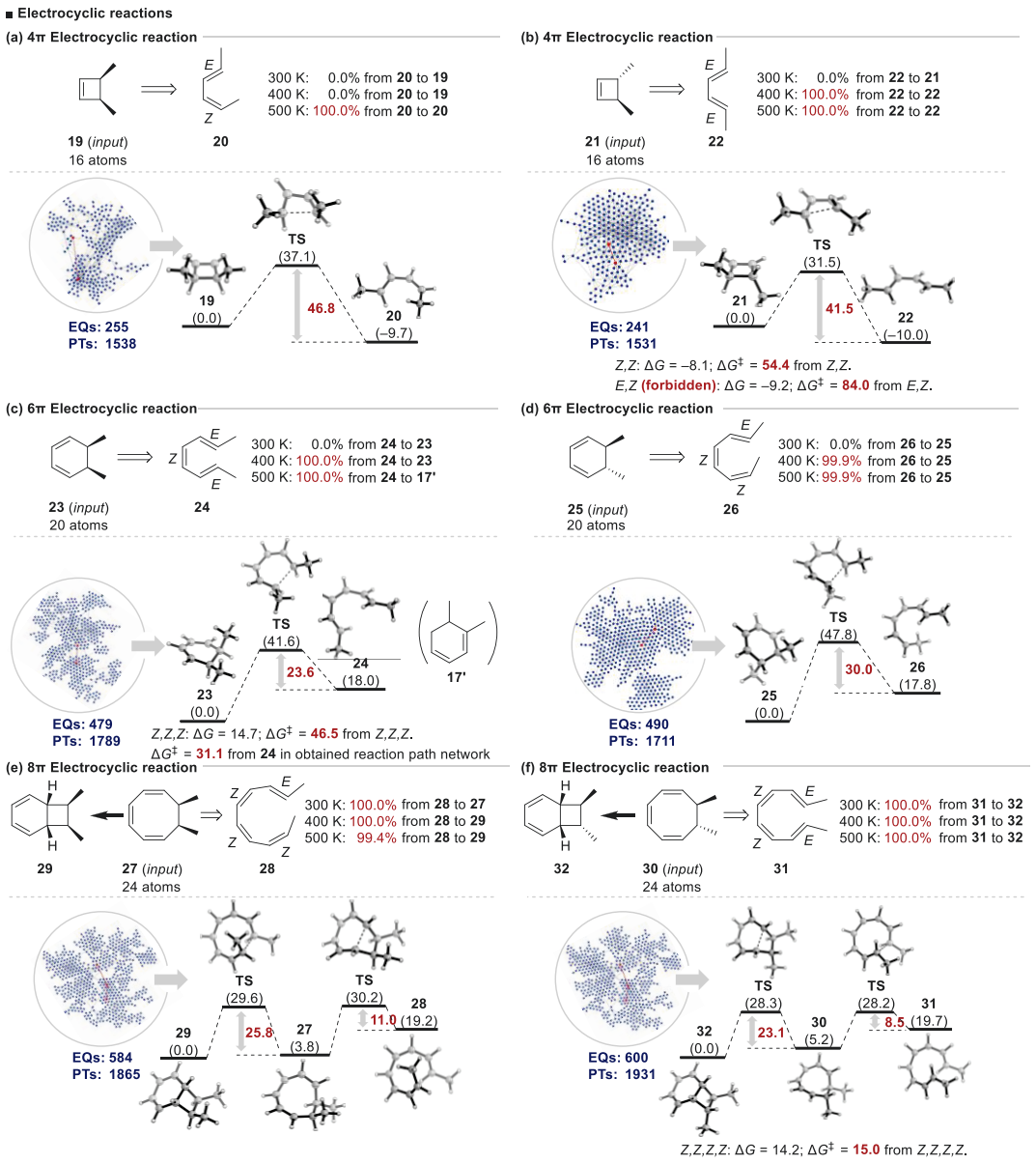


Figure 4. QCaRA for pericyclic reactions-2 (electrocyclic reactions). AFIR search: ωB97X-D/SV; structure refinement: ωB97X-D/6-311+G(d,p)//ωB97X-D/6-31G(d,p).

correlation between each EQ and PT, in which the expected Diels–Alder reaction of 1,3-butadiene (**2**) with ethylene (**3**) (red line in the network) exists. The N-yields based on the starting materials **2** and **3** were calculated to be 100% both at 400 and 500 K, but 0% at 300 K under the applied calculation conditions (maximum 1000 paths). It should be noted here that these N-yields are calculated using the reaction path network at a relatively low computational level (ωB97X-D/SV). The following structure refinement and single-point calculations at the ωB97X-D/6-31G(d,p) and ωB97X-D/6-311+G(d,p) levels, respectively, provided the real activation energy with $\Delta G^{\ddagger}$ = 28.2 kcal/mol, which is in good agreement with the reported value (27 kcal/mol).[4]

Using the same calculation conditions, 1-pentene (**4**) was subjected to the QCaRA (Figure 3b, 15 atoms) to demonstrate that products including propene (**5**) and ethylene (**3**) were successfully found via a retro-ene reaction. A total of 124 EQs and 1367 PTs were automatically generated in this search, and **4** was obtained in 100% N-yield at 400 K or higher from the proposed starting material pairs (**5** + **3**). The high yield at 400 K, which seems to conflict with the high activation barrier with 39.0 kcal/mol at the higher computational level, is due to the lower activation barrier with 24.2 kcal/mol at the lower computational level without the structure refinement on the reaction path network.

Next, we used benzene (**6**) and dihydrogen (**7**) as inputs (Figure 3c; 14 atoms).[23] 1,4-Cyclohexadiene (**8**) was obtained from group-transfer pericyclic reactions, while 1,3-cyclohexadiene (**9**) was also obtained with an extremely high activation barrier; a total of 315 EQs and 1613 PTs were automatically generated. However, both N-yields were 0% at all temperatures, most likely due to the prohibitively high activation barriers. The activation barrier for the path from **8** was 47.0 kcal/mol, while that from **9** was 119.2 kcal/mol. The former group-transfer reaction is [π2s + π2s + σ2s], which is the thermally allowed pathway. The forbidden [π2s + σ2s] pathways, leading to **9**, were not observed, albeit an alternative stepwise pathway wherein dihydrogen approaches twisted

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985–23000

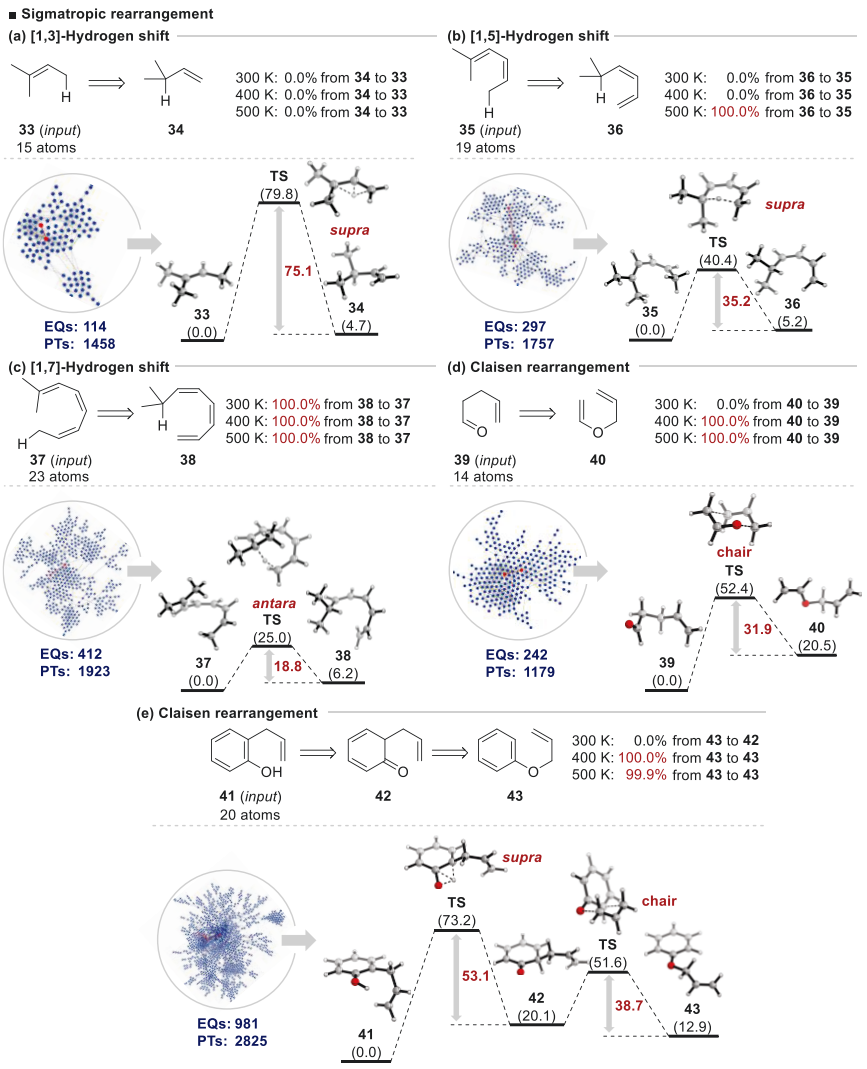


Figure 5. QCaRA for pericyclic reactions-3 (sigmatropic rearrangements). AFIR search: ωB97X-D/SV; structure refinement: ωB97X-D/6-311+G(d,p)//ωB97X-D/6-31G(d,p).

benzene and cleavage of one molecule of dihydrogen is followed by trapping with another molecule of dihydrogen through the shallow intermediate (**6′** + 7) (green line in the network). The latter pathway has an extremely high activation barrier so that attempted experiments using 1,3-cyclohexadine did not afford benzene and hydrogen.[23]

Cheletropic Reactions. Given the reasonable computational outcomes from the successful retrosynthetic analyses of three typical pericyclic reactions (Diels–Alder, ene, and group-transfer reactions), we subsequently focused on cheletropic reactions as our first targets for stereospecific transformations.

When the five-membered cyclic alkene **10** with the *cis*-orientation was subjected to the QCaRA, (*E,E*)-2,4-hexadiene (**11**) and sulfur dioxide (**12**) were formed via the thermally allowed cyclization mode (Figure 3d; 19 atoms).[24] (*E,E*)-isomer **11** was formed predominantly, while the formation of thermally forbidden (*E,Z*)-isomer **14** was not observed. In contrast, when cyclic alkene **13** with a *trans*-orientation was used instead of *cis*-isomer **10**, (*E,Z*)-2,4-hexadiene (**14**) was obtained exclusively (Figure 3e; 19 atoms). The number of calculated EQs and PTs was as follows: 914 EQs and 1811 PTs for **10** and 935 EQs and 1808 PTs for **13**. Given the relatively low activation barriers in both cases, N-yields of 100% were observed, which results in a complete equilibrium between the inputs and corresponding dienes with **12**. On the reaction path network, **13** further transforms into **13′** at 500 K. In other words, **13′** should be the actual product at 500 K. Therefore, the N-yield of 100% at 500 K refers to the production of **13′** as indicated in Figure 3e.

Subsequently, when seven-membered cyclic alkene **15** with a *cis*-orientation was subjected to the QCaRA, (*E,Z,Z*)-2,4,6-octatriene (**16**) and sulfur dioxide (**12**) were obtained in a stereospecific manner (Figure 3f; 23 atoms),[25] while the formation of (*E,Z,E*)-2,4,6-octatriene (**18**) was not observed. In contrast, when *trans*-isomer **17** was employed instead of *cis*-isomer **15**, (*E,Z,E*)-2,4,6-octatriene (**18**) was obtained predominantly (Figure 3g; 23 atoms). The observed thermally allowed antarafacial additions to sulfur dioxide (**12**) are fully consistent with the Woodward–Hoffmann rules. The number of calculated EQs and PTs was as follows: 997 EQs and 1883 PTs for **15** and 986 EQs and 1822 PTs for **17**. The N-yields were calculated as 100% at 300 and 400 K on account of the equilibrium between the inputs and triene with **12**. However, at 500 K, the N-yields are reduced to 57.8 and 42.7%, respectively, because more stable **15′** and **17′** have emerged and partly accessed from the reactant pairs. The cyclic

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

products **10**, **13**, **15**, and **17** are slightly more stable than polyenes **11**, **14**, **16**, and **18**, but the reaction preferentially proceeds in a desulfonylative fashion with the release of SO_2 gas because of its removal under experimental heating conditions.[24,25] These calculated results for cheletropic reactions indicate that the AFIR calculations can follow the stereospecific retrosynthesis analysis and that thermally forbidden pathways that have a high energy barrier are avoided automatically.

Electrocyclic Reactions. We next applied QCaRA to electrocyclic reactions, wherein the occurrence of a conrotatory or disrotatory process is governed by the number of π-electrons involved in the reactions (4π: conrotatory; 6π: disrotatory; 8π: conrotatory under thermal conditions).[5] Our first target was 4π electrocyclic reactions with methyl-disubstituted *cis*-cyclobutene **19** (Figure 4a; 16 atoms). The reactant bearing *E,Z* isomer **20** was generated exclusively via the expected conrotatory process in a stereospecific manner after the calculation of 255 EQs and 1538 PTs. Accordingly, its *trans*-diastereomer, that is, *trans*-3,4-dimethylcyclobutene (**21**), was examined to demonstrate that the expected conrotatory process proceeded selectively to afford both *E,E* isomer **22** and the corresponding *Z,Z* isomer (Figure 4b; 16 atoms) after the calculation of 241 EQs and 1531 PTs. The N-yield at 300 K was 0% because **19** and **21** cannot be accessed from **20** and **22**, respectively, because of prohibitively high activation barriers. At higher temperatures (especially 500 K), **20** and **22** were recognized as the most stable compounds on each reaction map and their N-yields inevitably turned to 100%. In the case of **21**, the forbidden disrotatory process was also observed, albeit its activation barrier is very high ($\Delta G^{\ddagger}$ = 84.0 kcal/mol; calculated for the closed-shell singlet excluding the possibility of a diradical mechanism).

We next investigated the 6π electrocyclic reaction of *cis*-5,6-dimethyl-1,3-cyclohexadiene (**23**) (Figure 4c; 20 atoms). Only the expected disrotatory processes were observed after the calculation of 492 EQs and 1789 PTs to afford both *E,Z,E* isomer **24** and the corresponding *Z,Z,Z* isomer. The $\Delta G^{\ddagger}$ from the *Z,Z,Z* isomer is significantly higher than that from *E,Z,E* isomer **24** ($\Delta G^{\ddagger}$ = 46.5 vs 23.6 kcal/mol). The N-yield at 300 K was 0% because the activation barrier in the reaction path network calculated at the ωB97X-D/SV level was slightly higher (31.1 kcal/mol). Moreover, the N-yield at 500 K was calculated to be 100%, because more stable **17′** can be accessed from **23** through a [1,5]-hydrogen shift. In contrast, when *trans*-5,6-dimethyl-1,3-cyclohexadiene (**25**) was employed in the calculations, only the formation of *E,Z,Z* isomer **26** was observed after the calculation of 490 EQs and 1711 PTs (Figure 4d; 20 atoms).

Moreover, 8π electrocyclic reactions were examined using *cis*-7,8-dimethyl-1,3,5-cyclooctatriene (**27**). After the calculation of 584 EQs and 1865 PTs, the formation of the *E,Z,Z,E* isomer **28** was observed together with the energetically more stable bicyclic compound **29**, both via thermally allowed electrocyclic reaction processes (8e: conrotatory; 6e: disrotatory) (Figure 4e; 24 atoms). When *trans*-isomer **30** was subjected to the QCaRA, the formation of *E,Z,Z,E* isomer **31** was observed after the calculation of 600 EQs and 1931 PTs, together with the energetically more stable bicyclic product **32**, both of which were generated via thermally allowed processes (8e: conrotatory; 6e: disrotatory) (Figure 4f; 24 atoms). The N-yield of *cis*-isomer **27** from **28** at 300 K was calculated to be 100%; at 400 and 500 K, the activation barrier of the ring-closing 6π electrocyclic reaction can be overcome, which resulted in the formation of the more stable **29** in 100% N-yield. On the other hand, in the case of *trans*-isomer **30**, all N-yields were 100% from **31** to **32**.

Sigmatropic Rearrangements. Our next target was [*m*,*n*]-sigmatropic rearrangements with a variety of reaction modes, including [1,3]-, [1,5]-, and [1,7]-hydrogen shifts as well as [3,3]-Claisen and Cope rearrangements. We first subjected 2-methyl-2-butene (**33**) to the QCaRA calculations (Figure 5a; 15 atoms). The [1,3]-hydrogen shift is generally regarded as an antarafacial hydrogen shift under thermal conditions, albeit this antarafacial process is strictly forbidden because of the sterically inaccessible transfer. The AFIR proposed 114 EQs and 1458 PTs, which included the suprafacial forbidden shift with a very high activation barrier (75.1 kcal/mol). No pathways were observed for the thermally allowed but sterically forbidden antarafacial hydrogen shift. When 2-methyl 2,4-hexadiene (**35**) was subjected to the QCaRA, 297 EQs and 1757 PTs were obtained, which included the thermally allowed suprafacial [1,5]-hydrogen shift (Figure 5b; 19 atoms), while the thermally forbidden antarafacial hydrogen shift was not observed. Furthermore, when 2-methyl-2,4,6-octatriene (**37**) was employed in the calculations, the antarafacial [1,7]-hydrogen shift was observed exclusively after the calculation of 412 EQs and 1923 PTs (Figure 5c; 23 atoms). Notably, these stereospecific trends are fully consistent with the Woodward−Hoffmann rules.

Next, we investigated a [3,3]-sigmatropic reaction (Claisen rearrangement). When 4-pentenal (**39**) was subjected to the QCaRA, the corresponding allylvinylether (**40**) was produced via a Claisen rearrangement after the calculation of 242 EQs and 1179 PTs with a reasonable $\Delta G^{\ddagger}$ (31.9 kcal/mol) (Figure 5d; 14 atoms).[26] We then investigated the more challenging substrate 2-allylphenol (**41**) because it has a high possibility to be retrosynthesized into allylphenylether using the QCaRA (Figure 5e; 20 atoms).[27] The initial keto−enol tautomerization was observed, albeit this step exhibits a very high energy barrier ($\Delta G^{\ddagger}$ = 53.1 kcal/mol) because this process is categorized as a thermally forbidden suprafacial [1,3]-hydrogen shift. The kinetic navigation system (RCMC) judged that the generation of keto product **42** is energetically unfavorable, even at higher temperatures such as 500 K;[20] thus, the subsequent path search from keto product **42** was not conducted within the 1000-path search. Therefore, the calculation of an additional 1000 paths from keto product **42** was carried out to successfully afford the reactant allylphenylether (**43**) with a reasonable $\Delta G^{\ddagger}$. A total of 981 EQs and 2825 PTs were produced in this series of two calculations.

Keto−enol tautomerizations are known to be facile processes, wherein two or more molecules are involved in a real experimental system. However, QCaRA is conducted using only one input molecule without any protic or basic additives, which usually results in a significant difference between the calculated and the experimentally observed results. Indeed, when keto product **42** and a proton (H^+) were used as inputs, 2-allylphenol (**41**) was quickly generated via a barrierless protonation. With respect to N-yields, the reference compound for the determination of N-yields using RCMC was **42** at 300 K (0%), but turned to **43** at 400 and 500 K (100.0 and 99.9%, respectively) because **41** is not accessible due to the extremely high activation barrier of thermally forbidden keto-enol tautomerization. These results clearly indicate that the automatically generated reaction pathways

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

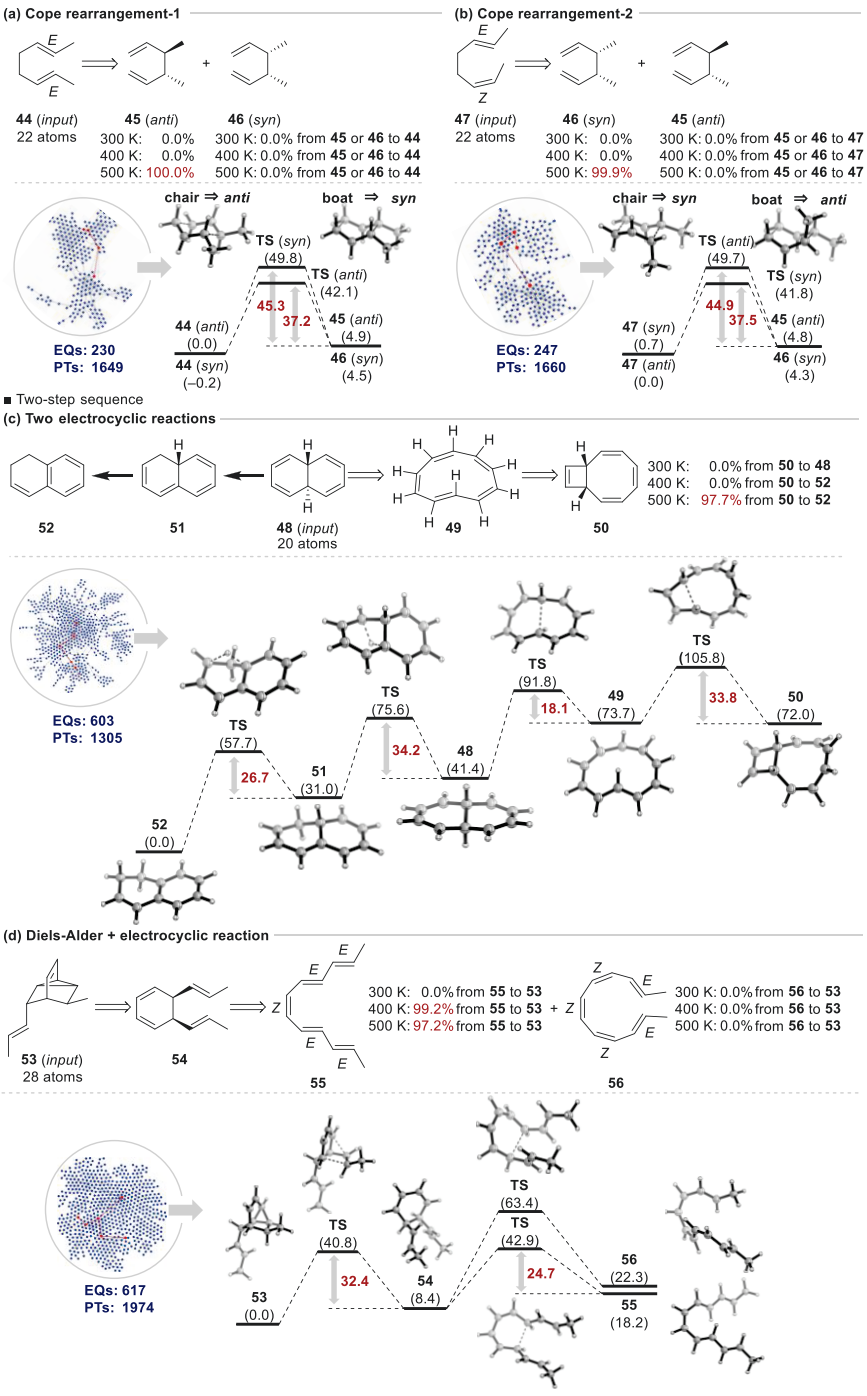


Figure 6. QCaRA for pericyclic reactions-4 (sigmatropic rearrangements and two-step sequences). AFIR search: ωB97X-D/SV; structure refinement: ωB97X-D/6-311+G(d,p)//ωB97X-D/6-31G(d,p).

involved in the keto−enol tautomerization differ from real experimental systems and that the precise reaction diagram cannot be predicted if a molecule that catalyzes the process is omitted.

Our next target reaction was the stereoselective Cope rearrangement. We first examined (*E*,*E*)-2,6-octadiene (**44**) in the QCaRA calculations (Figure 6a; 22 atoms). After the calculation of 230 EQs and 1649 PTs, the formation of both *anti*- and *syn*-products **45** and **46** was observed; the former came from the more stable chair TS, while the latter comes from the less stable boat TS. The activation barrier of the chair TS to afford the *anti*-product is lower than that of the boat TS to afford the *syn*-product, all of which is fully consistent with the experimental results.[28] In contrast, when (*E*,*Z*)-2,6-octadiene (**47**) was used in the calculations, the formation of both the *syn*- and *anti*-products **46** and **45** was observed after the calculations of 247 EQs and 1660 PTs, respectively; **46** is derived from the more stable chair TS, while **45** is derived from the less stable boat TS (Figure 6b; 22 atoms). The activation barrier of the chair TS to afford the *syn*-product is lower than that of the boat TS to afford the *anti*-product. At 500 K, N-yields of 100 and 99.9% were observed from *anti*-isomer **45** to (*E*, *E*)-isomer **44**, and from *syn*-isomer **46** to (*E*, *Z*)-isomer **47**, respectively, which is all in accordance with the

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

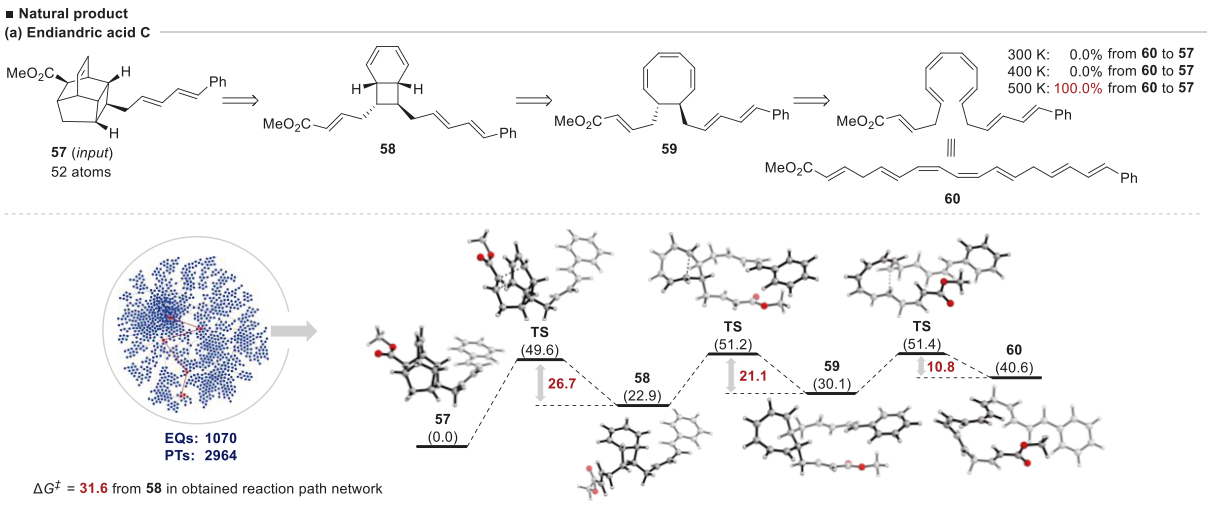


Figure 7. QCaRA for pericyclic reactions-5 (natural products). AFIR search: ωB97X-D/SV; structure refinement: ωB97X-D/6-311+G(d,p)//ωB97X-D/6-31G(d,p).

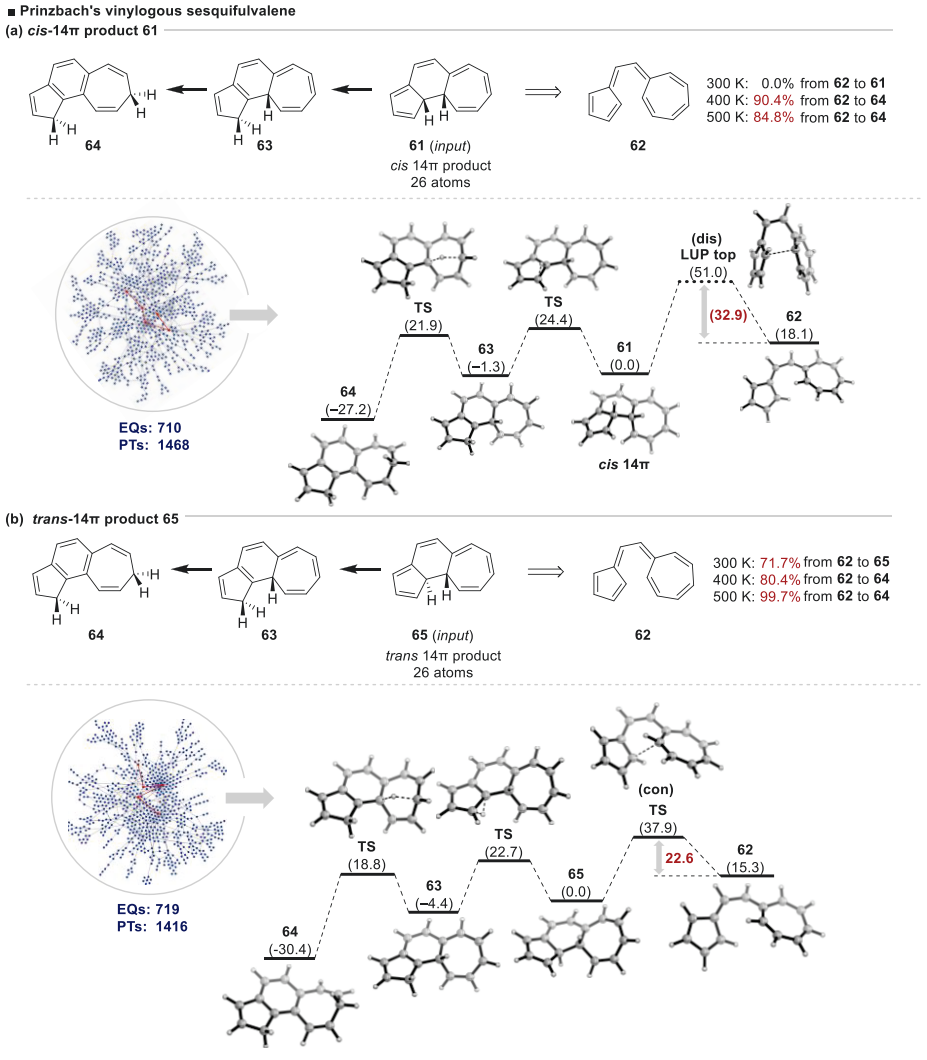


Figure 8. QCaRA for pericyclic reactions-6 (cyclized products for Prinzbach's vinylogous sesquifulvalene). AFIR search: ωB97X-D/SV; structure refinement: ωB97X-D/6-311+G(d,p)//ωB97X-D/6-31G(d,p).

expected selectivity of Cope rearrangements. Given that the observed selectivity is determined by the geometry of the TSs, the formation of the more stable chair TSs is ca. 8 kcal/mol lower than that of unstable boat TSs,[29] both of which can now be automatically searched using the AFIR. It should also be noted here that all calculated energy values in the

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

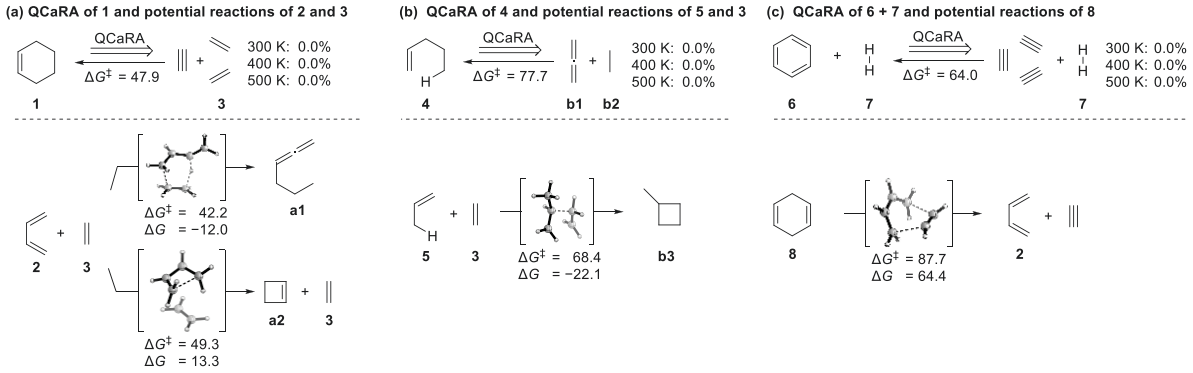


Figure 9. Other potential reactions extracted from the reaction path network-1. AFIR search: ωB97X-D/SV; without structure refinement.

aforementioned standard pericyclic reactions are very reasonable in comparison with the calculated values reported by Houk.[4]

Two-Step Sequences. We then turned our attention to a more complex system that involves two sequential electrocyclic reactions (Figure 6c; 20 atoms). When *trans*-dihydronaphthalene (**48**) was subjected to the QCaRA, the initial 6π electrocyclic reaction occurred in a disrotatory manner to afford ten-membered polyene **49**, which includes an *E*-alkene moiety. A further conrotatory 4π electrocyclic reaction[1,2] gave *cis*-conjugated product **50**. Those two reaction pathways are all thermally allowed. In addition, the more stable aromatic compound **52** was observed via two suprafacial [1,5]-hydrogen shifts. A total of 603 EQs and 1305 PTs were generated in this search. The calculated N-yields suggested that **48** would probably not be obtained from **50** at 300 and 400 K. However, at 500 K, two consecutive [1,5]-hydrogen shifts can be expected to furnish the relatively stable aromatic compound **52**.

Next, the rather complex molecule **53** was subjected to the QCaRA (Figure 6d; 28 atoms). Initially, 5,6-disubstituted 1,3-cyclohexadiene **54** was formed via a retro-Diels–Alder reaction and was further retrosynthesized into *E,E,Z,E,E* polyene **55** and *E,Z,Z,Z,E* polyene **56** via a disrotatory 6π electrocyclic reaction in both cases.[30] The activation energy for the generation of sterically more crowded **56** is much higher than that of **55**. A total of 617 EQs and 1974 PTs were generated in this case. Thus, molecules with over 20 atoms, such as **48** and **53**, were successfully retrosynthesized using the QCaRA to afford the reactants in two steps via thermally allowed processes without using any empirical templates. At 400 and 500 K, the N-yield from **55** to **53** is almost 100%.

Natural Products. Our ultimate QCaRA target as the most complicated molecule was the natural product endiandric acid C (**57**), whose methyl ester was chemically synthesized by Nicolaou in 1982[14] via three pericyclic reactions in one pot at 100 °C based on Black's biosynthetic hypothesis for endiandric-acid analogues.[15] Even though this molecule contains more than 50 atoms (Figure 7a; 52 atoms), the AFIR could manage its retrosynthesis very well within the calculation of 1000 paths including a total of 1070 EQs and 2964 PTs. Bicyclic compound **58** was first generated via a retro-Diels–Alder reaction followed by a disrotatory 6π electrocyclic reaction to afford **59** with a *trans*-orientation. The final ring-opening proceeded via a conrotatory 8π electrocyclic reaction to give the starting polyene **60** with the correct alkene geometry, which is 40.6 kcal/mol higher in energy than **57**.[31] Nicolaou's experimental result that **58** is formed exclusively before heating to 100 °C[14d] is clearly explained by the obtained diagram, in which the $\Delta G^{\ddagger}$ between **58** and **57** is 26.7 kcal/mol. At 400 K (123 °C), an N-yield of 0% was observed, which is inconsistent with Nicolaou's experiment given that the activation barrier between **58** and **57** in the reaction path network at the ωB97X-D/SV level was slightly higher (31.6 kcal/mol) than that after the structure refinement and single-point calculations (26.7 kcal/mol). However, notably, the AFIR successfully retrosynthesized a molecule containing 52 atoms to locate higher-energy compounds with the proper stereochemistry, which represents a significant advancement of automated reaction path search methods for the design of chemical reactions.

Prinzbach's Vinylogous Sesquifulvalene. The final example is a case study on a reaction that violates the Woodward–Hoffmann rules. Very recently, Houk and coworkers reported an elegant mechanistic study of thermally forbidden 14π electrocyclic reactions for Prinzbach's vinylogous sesquifulvalene **62**,[32] wherein the thermically allowed disrotatory 14π electrocyclic reaction that affords *cis*-product **61** is disfavored compared to the thermally forbidden conrotatory pathway that affords *trans*-product **65** because of the electronic and steric constraints in the TS (violations of the Woodward–Hoffmann rules).[16] Encouraged by Houk's report, we first subjected the disfavored *cis*-product **61** to the QCaRA (Figure 8a; 26 atoms), which successfully afforded reactant **62** after the calculation of 710 EQs and 1468 PTs. In addition, the formation of the more stable indene **64** was observed after two suprafacial [1,5]-hydrogen shifts. The obtained N-yield corroborates the unfavorableness of the thermally allowed disrotatory electrocyclic reaction. At 300 K, the reaction of **62** cannot proceed because of its high activation barrier, while at 400 and 500 K, **64** was obtained through energetically lower intermediate *trans*-product **65** (*vide infra*). Unfortunately, we could not determine the exact TS between **61** and **62** as this TS easily converged to a different TS between *cis*-product **61** and the corresponding 10π product[16] during TS optimization, and the TS was approximated by the highest energy point along the path optimized by the LUP method. Even with the approximate TS, however, the activation energy ($\Delta G^{\ddagger}$ = 32.9 kcal/mol) is very similar to the value calculated by Houk ($\Delta G^{\ddagger}$ = 33.5 kcal/mol).[16]

Next, the favored *trans*-product **65** was calculated, which generated reactant **62** and indene **64** after 719 EQs and 1416 PTs (Figure 8b; 26 atoms). The N-yield of *trans*-product **65** indicated a different tendency from that of the *cis*-product: At 300 K, **65** would be obtained as the product from **62**, while at 400 and 500 K, the product would be altered from **62** to **64**.

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

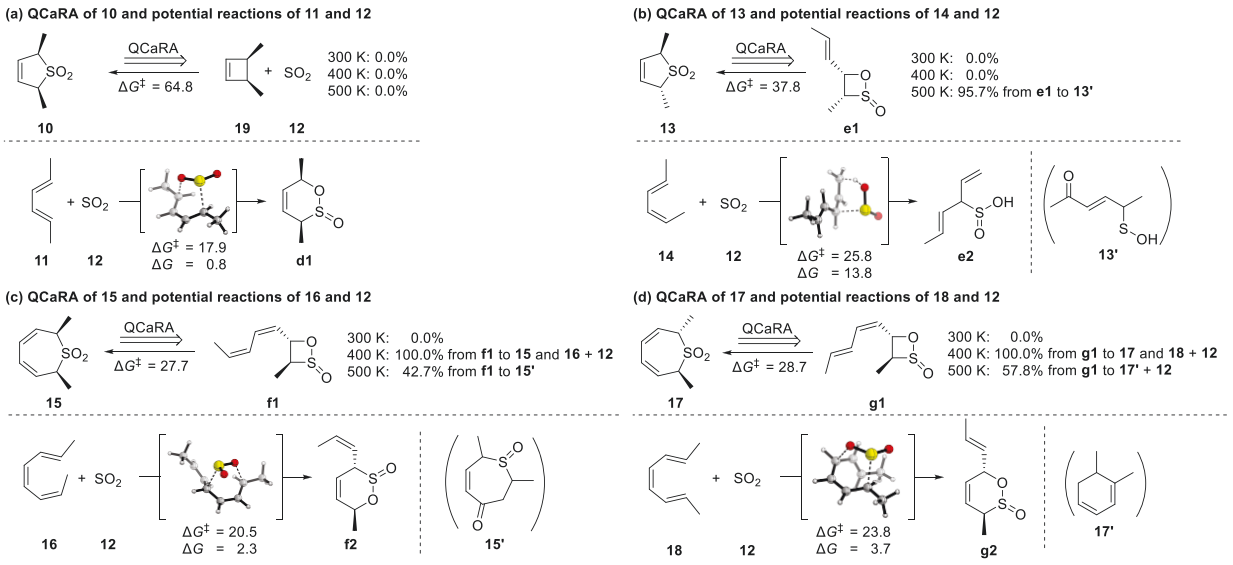


Figure 10. Other potential reactions extracted from the reaction path network-2. AFIR search: ωB97X-D/Def2-SV(P) for S atoms and SV for other atoms; without structure refinement.

The thermally forbidden, but experimentally accessible, TS between **62** and **65** was calculated to be $\Delta G^{\ddagger}$ = 22.6 kcal/mol, which is very similar to the value calculated by Houk ($\Delta G^{\ddagger}$ = 22.3 kcal/mol). In addition, the formation of the more stable indene **64** was also observed after two suprafacial [1,5]-hydrogen shifts, which is in good agreement with the experimental result by Prinzbach.[16] It should also be noted here that the thermally forbidden, but experimentally accessible pathway can be searched automatically, which greatly enhances the practical utility of this AFIR-based retrosynthetic analysis. Reported pericyclic reactions are usually conducted under neat conditions[23−26,28] or using a nonpolar solvent (toluene or benzene)[14d,27,30,32] at elevated temperatures, which is well suited to our gas-phase calculations.

Other Potential Reactions. As we created a reaction path network for each type of representative pericyclic reactions, more details on the network of representative pericyclic reactions (excluding two-step sequences and natural products) will be scrutinized and discussed in this section, especially other potential retrosynthetic pathways as well as the feasibility of the obtained starting-material pairs. The amenable retrosynthetic pathways (generally the second lowest TS, excluding minor bond rotation and unrealistic carbene formation) and forward pathways from the proposed starting-material pairs with low activation barriers (generally the second lowest TS, excluding minor bond rotation and unrealistic carbene formation) are shown in Figures 9−13. The energy values shown in Figures 9−13 are directly derived from the values of the reaction maps using PTs (rough TS) at the ωB97X-D/SV level without structure refinement and single-point calculations at the higher computational levels.

The QCaRA calculations of cyclohexene (**1**) provided not only the Diels-Alder reaction route in Figure 3a but also another possible retrosynthetic pathway to two molecules of ethylene (**3**) and acetylene via [2 + 2 + 2] cyclization (Figure 9a). However, this pathway has a high activation barrier ($\Delta G^{\ddagger}$ = 47.9 kcal/mol) that cannot be overcome in the examined temperature window (0% at 300, 400, and 500 K). In the forward pathway, the reaction of proposed **2** and **3** computationally afforded 1-propylallene (**a1**) and cyclobutene (**a2**) via an ene-type reaction and 4π conrotatory electrocyclic reaction, respectively, albeit the activation barrier for both pathways are very high ($\Delta G^{\ddagger}$ = 42.2 and 49.3 kcal/mol, respectively). These results clearly indicate that **2** and **3** are the energetically most accessible compounds in the retrosynthetic analysis of **1** and the pathway to **1** from **2** and **3** should be the most accessible among the calculated pathways. Next, the QCaRA of 1-pentene (**4**) afforded not only **5** and **3** but also allene (**b1**) and ethane (**b2**) via the insertion of the C−H bond of ethane to the double bond of allene with a very high activation barrier ($\Delta G^{\ddagger}$ = 77.7 kcal/mol) (Figure 9b). In the forward pathway, a thermally forbidden intermolecular [2 + 2] cyclization afforded methylcyclobutane (**b3**), albeit the associated energy barrier is extremely high ($\Delta G^{\ddagger}$ = 68.4 kcal/mol). The QCaRA of **6** and **7** provided three molecules of acetylene and **7** with $\Delta G^{\ddagger}$ = 64.0 kcal/mol (Figure 9c). The pathways from proposed **8** include a retro Diels-Alder reaction that affords the energetically higher-lying 1,3-butadiene (**2**) and acetylene (endergonic transformation), which is difficult to access experimentally.

As for cheletropic reactions, the QCaRA of **10** provided not only **11** + **12** but also *cis*-1,2-dimethylcyclobutene (**19**) with a very high activation energy ($\Delta G^{\ddagger}$ = 64.8 kcal/mol). Interestingly, the proposed pair **11** + **12** underwent the Diels-Alder-type [4 + 2] cycloaddition to afford slightly endergonic *cis*-oriented six-membered sulfinate **d1** with a reasonably low activation barrier so that **d1**, **10**, and **11** + **12** are all in equilibrium, even at 300 K (*vide supra*) (Figure 10a). Vogel has demonstrated that **11** reacts reversibly with **12** to afford **d1** at low temperatures. However, above −50 °C, **d1** undergoes fast retro-cyclization liberating starting materials **11** and **12**, which is consistent with these calculation results.[33] The QCaRA of **13** provided four-membered cyclic sulfinate **e1** with $\Delta G^{\ddagger}$ = 37.8 kcal/mol (Figure 10b). At 500 K, cyclic sulfinate **e1** could be transformed to the most stable compound **13′**, which can be energetically accessed in the obtained reaction map. In the forward pathways, the reaction of **14** + **12** provided 3-sulfinyl-1,4-diene **e2** via a highly endergonic pathway with $\Delta G^{\ddagger}$ = 25.8 kcal/mol. The QCaRAs

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

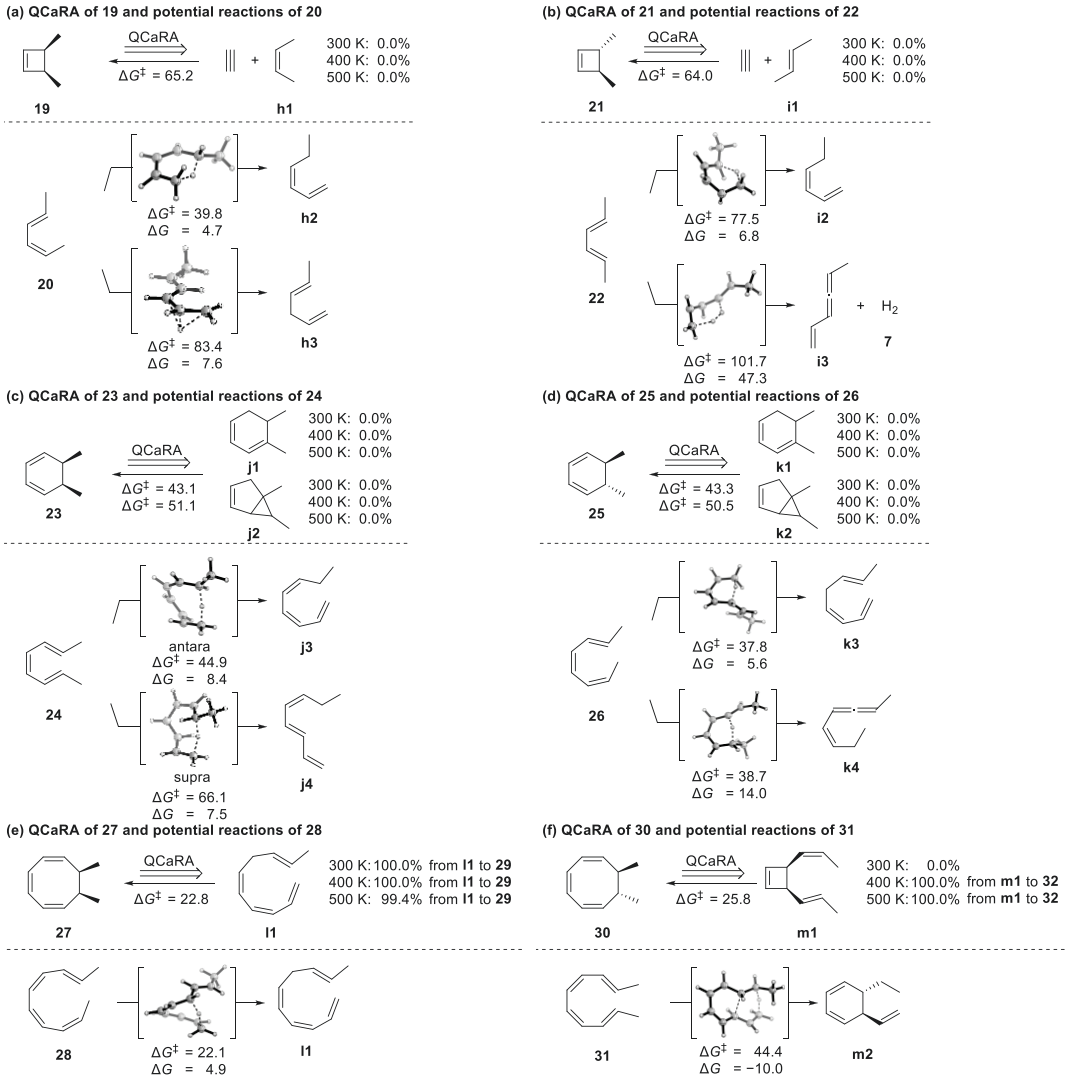


Figure 11. Other potential reactions extracted from the reaction path network-3. AFIR search: ωB97X-D/SV; without structure refinement.

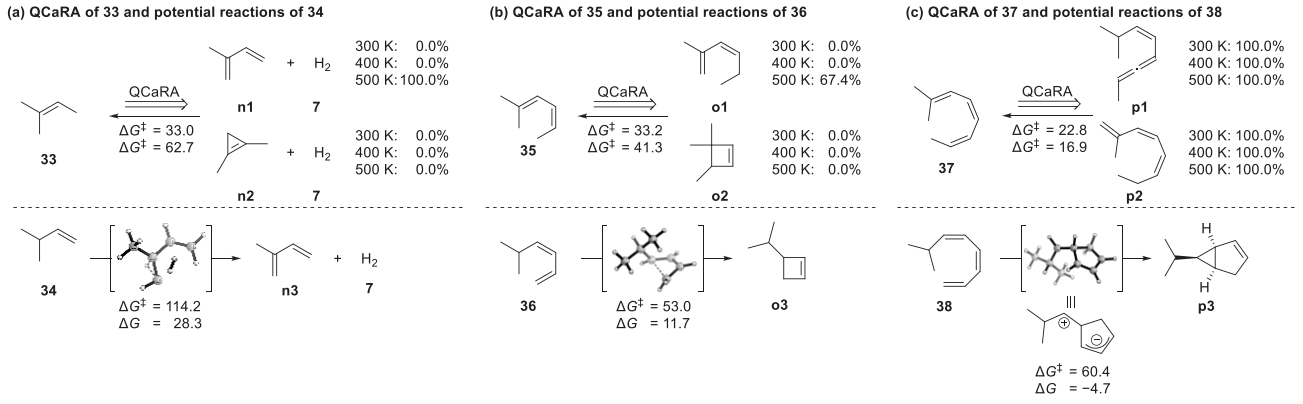


Figure 12. Other potential reactions extracted from the reaction path network-4. AFIR search: ωB97X-D/SV; without structure refinement.

of **15** or **17** afforded four-membered sulfinates **f1** or **g1**, which differ with respect to their olefin geometry. At 400 K, **f1** or **g1** can be expected to be in equilibrium with **15** and **16** + **12** or with **17** and **18** + **12**. On the other hand, the proposed starting material pairs **16** + **12** and **18** + **12** provided six-membered cyclic sulfinates **f2** or **g2** via slightly endergonic pathways (Figure 10c,d).

With respect to electrocyclic reactions, the QCaRA of **19** provided not only **20** but also acetylene and *Z*-oriented 2-butene **h1** via thermally forbidden [2 + 2] cyclization with a high activation barrier (Figure 11a). The proposed starting material **20** afforded mainly two alkene isomers (**h2** and **h3**) via a thermally allowed suprafacial [1,5]-hydrogen shift and a thermally forbidden suprafacial [1,3]-hydrogen shift, whereby

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

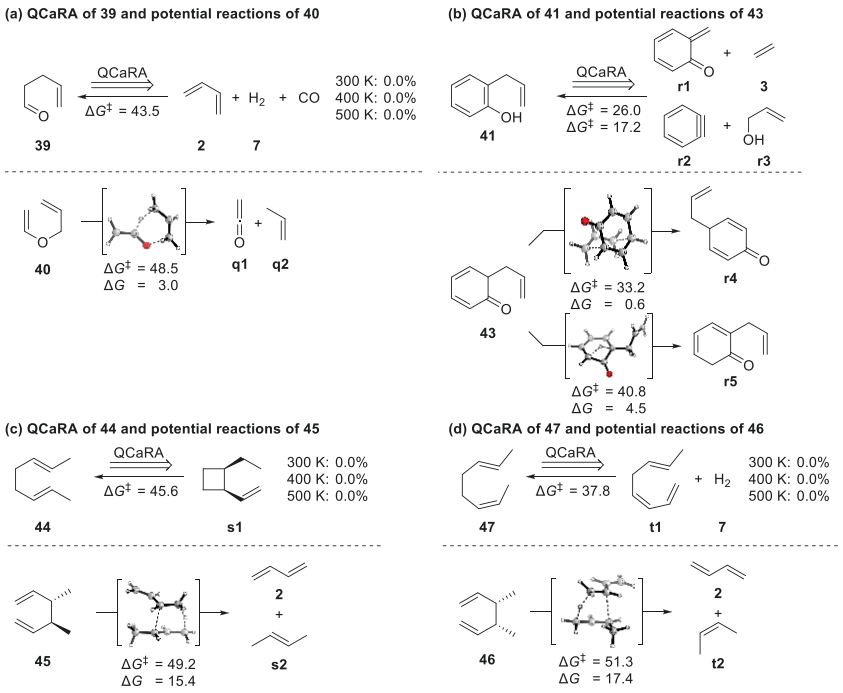


Figure 13. Other potential reactions extracted from the reaction path network-5. AFIR search: ωB97X-D/SV; without structure refinement.

the latter has an extremely high activation barrier. The QCaRA of **21** provided not only **22** but also acetylene and *E*-oriented 2-butene **i1** via a thermally forbidden [2 + 2] cyclization with a high activation barrier (Figure 11b). The proposed starting material **22** afforded the corresponding alkene isomer **i2** and allenyl alkene **i3** via a thermally allowed [1,5]-hydrogen shift and dehydrogenation, respectively.

The QCaRA of **23** provided both cyclohexadiene **j1** and cyclopentene **j2** with $\Delta G^{\ddagger}$ = 40−50 kcal/mol (Figure 11c). Moreover, the reaction of **24** provided **j3** and **j4** via a thermally allowed antarafacial [1,7]-hydrogen shift and a suprafacial [1,7]-hydrogen shift with an *E*-alkene, both of which obey the Woodward−Hoffmann rules. Furthermore, the QCaRA of **25** provided cyclohexadiene **k1** and cyclopentene **k2** with $\Delta G^{\ddagger}$ = 40−50 kcal/mol (Figure 11d). The QCaRA of proposed **26** provided **k3** and **k4** via thermally allowed suprafacial [1,5]-hydrogen shifts. Both activation barriers are relatively high (~40 kcal/mol). In addition, the QCaRA of **27** provided linear decatetraene **l1** with a reasonably low activation barrier (Figure 11e). At all temperatures, the energetically most stable bicyclic compound **29** (*vide supra*) should be obtained from **l1** via **27**. In the forward pathways, the reaction of **28** proceeded via a thermally allowed antarafacial [1,7]-hydrogen shift to afford **l1**. The QCaRA of *trans*-isomer **30** provided 1,2-*cis*-disubstituted cyclobutene **m1** via a Cope rearrangement (Figure 11f). At 400 and 500 K, the energetically most stable bicyclic compound **32** (*vide supra*) can be obtained from **m1** via **30**. In the forward pathways, the QCaRA of **31** provided 1,3-cyclohexadiene **m2** via an intramolecular ene-reaction with $\Delta G^{\ddagger}$ = 44.4 kcal/mol.

As for [1,n]-hydrogen shifts, the QCaRA of **33** provided 1,3-diene **n1** with H_2 and cyclopentene **n2** with H_2 as starting materials (Figure 12a). In both cases, the activation barriers are too high to be overcome at 300 and 400 K. The proposed starting material **34** could lead to 1,3-diene **n3** and H_2 via a dehydrogenation with an extremely high activation barrier. Next, the QCaRA of **35** proposed two starting material candidates, that is, a 1,3-diene bearing the *exo* olefin (**o1**) and cyclobutene **o2** (Figure 12b). The forward pathways from **36** led to cyclobutene **o3** via a thermally allowed conrotatory 4π electrocyclic reaction, albeit this pathway is endergonic. The QCaRA of **37** provided **p1** and **p2** via reasonable activation barriers and the N-yield from each compound was 100% at all temperatures, albeit these compounds seem to be very difficult to prepare experimentally (Figure 12c). In the forward pathway, the proposed starting material **38** led to the formation of bicyclic product **p3** with a very high activation barrier ($\Delta G^{\ddagger}$ = 60.4 kcal/mol) via an ionic transition state.

Finally, we will discuss reaction path networks of several sigmatropic rearrangements. The QCaRA of **39** provided not only **40** but also 1,3-butadiene (**2**), H_2 (**7**), and CO in a noncatalytic retro-hydroformylation manner with $\Delta G^{\ddagger}$ = 43.5 kcal/mol (Figure 13a). In the forward pathway, the proposed **40** led to ketene (**q1**) and 1-propene (**q2**) via a retro-ene reaction. The QCaRA of **41** provided enone **r1** and ethylene (**3**) as well as unstable benzyne (**r2**) and allylic alcohol (**r3**) (Figure 13b). In the forward pathway, **43** underwent a Cope rearrangement via the chair TS to give γ-allyl dienone **r4**. Enone **43** also undergoes a [1,5]-hydrogen shift to afford the corresponding structurally different enone **r5**, both of which are characterized by high activation barriers. In the QCaRA toward Cope rearrangements, **44** furnished the corresponding cyclobutane **s1** via an intramolecular ene-type reaction with $\Delta G^{\ddagger}$ = 45.6 kcal/mol (Figure 13c). In the forward pathways from **45**, 1,3-butadiene (**2**) and *trans*-2-butene (**s2**) were obtained with $\Delta G^{\ddagger}$ = 49.2 kcal/mol via an endergonic intramolecular ene-type reaction. The QCaRA of **47** afforded the corresponding triene (**t1**) and H_2 (**7**) via a group-transfer fragmentation ($\Delta G^{\ddagger}$ = 37.8 kcal/mol) (Figure 13d). In the forward reaction pathways, **46** engages in an intramolecular ene-type reaction to give 1,3-butadiene (**2**) and *cis*-2-butene (**t2**).

All the results described in this section clearly indicate that the most accessible and realistic pathways using readily

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

available starting materials affording input compounds should be the pericyclic reactions described in Figures 3−8 and that the most reasonable route from the proposed starting materials should be the expected pericyclic reactions as well.

CONCLUSIONS

In summary, we have conducted DFT-based automated reaction-path searches applying the QCaRA to several pericyclic reactions under unified conditions. The products arising from cycloaddition, ene, group-transfer, cheletropic, and electrocyclic reactions as well as sigmatropic rearrangements were all reasonably retrosynthesized to afford the reported starting materials in a stereospecific manner. In essence, this is a very useful method for identifying energetically higher-lying reactants from target molecules without any a priori experimental information or human intuition or expertise, instead using only quantum-chemical calculations that are guided by AFIR as a type of chemical compass. Even a natural product containing over 50 atoms as well as a product of a conrotatory 14π electrocyclic reaction, which violates the Woodward−Hoffmann rules, was successfully retrosynthesized with the correct stereochemistry. We also confirmed that the proposed reactant candidates are the most reasonable starting materials. This technology can be expected to open new avenues in the computational design of new chemical reactions from scratch. Efforts toward this goal are now being extensively examined in our laboratories, and the results will be disclosed elsewhere in due course.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.2c09830.

Computational methods, computational results by comparison of the γ value and basis sets, as well as cartesian coordinates of each EQ and TS (PDF)

AUTHOR INFORMATION

Corresponding Authors

Tsuyoshi Mita − *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo, Hokkaido 001-0021, Japan; JST, ERATO Maeda Artificial Intelligence in Chemical Reaction Design and Discovery Project, Sapporo, Hokkaido 060-0810, Japan;* orcid.org/0000-0002-6655-3439; Email: tmita@icredd.hokudai.ac.jp

Satoshi Maeda − *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo, Hokkaido 001-0021, Japan; JST, ERATO Maeda Artificial Intelligence in Chemical Reaction Design and Discovery Project, Sapporo, Hokkaido 060-0810, Japan; Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan; Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS), Tsukuba, Ibaraki 305-0044, Japan;* orcid.org/0000-0001-8822-1147; Email: smaeda@eis.hokudai.ac.jp

Authors

Hideaki Takano − *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo, Hokkaido 001-0021, Japan; JST, ERATO Maeda Artificial Intelligence in Chemical Reaction Design and Discovery Project, Sapporo, Hokkaido 060-0810, Japan;* orcid.org/0000-0003-0744-9292

Hiroki Hayashi − *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo, Hokkaido 001-0021, Japan; JST, ERATO Maeda Artificial Intelligence in Chemical Reaction Design and Discovery Project, Sapporo, Hokkaido 060-0810, Japan;* orcid.org/0000-0003-2081-6886

Wataru Kanna − *Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan;* orcid.org/0000-0002-2931-6661

Yu Harabuchi − *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo, Hokkaido 001-0021, Japan; JST, ERATO Maeda Artificial Intelligence in Chemical Reaction Design and Discovery Project, Sapporo, Hokkaido 060-0810, Japan; Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan;* orcid.org/0000-0001-8313-3236

K. N. Houk − *Department of Chemical and Biomolecular Engineering and Department of Chemistry and Biochemistry, University of California Los Angeles, Los Angeles, California 90095, United States;* orcid.org/0000-0002-8387-5261

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.2c09830

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was financially supported by JST-ERATO (JPMJER1903), JSPS-WPI, and Grants-in-Aid for Challenging Research (Exploratory) (21K18945), Scientific Research (B) (22H02069), Transformative Research Areas (A) (Digitalization-driven Transformative Organic Synthesis (Digi-TOS)) (22H05330), and Young Scientists (22K14673). T.M. thanks the Fugaku Trust for Medical Research, the Uehara Memorial Foundation, and the Naito Foundation for financial support. K.N.H. is grateful to the US National Science Foundation (CHE-1764328). All calculations were performed using the supercomputer Fugaku provided by the RIKEN Center for Computational Science as well as the supercomputer system at the Information Initiative Center, Hokkaido University, Sapporo, Japan. The use of the GRRM and Gaussian programs on the supercomputer Fugaku was enabled with the support of HPC Systems, Inc.

REFERENCES

(1) Fleming, I. *Pericyclic Reactions*, 2nd Edition; Oxford University Press: Oxford, UK, 2015.

(2) Woodward, R. B.; Hoffmann, R. The conservation of orbital symmetry. *Angew. Chem., Int. Ed.* **1969**, *8*, 781−853.

(3) (a) Zimmerman, H. E. On molecular orbital correlation diagrams, the occurrence of Möbius systems in cyclization reactions, and factors controlling ground- and excited-state reactions. *J. Am. Chem. Soc.* **1966**, *88*, 1564−1565. (b) Dewer, M. J. S. Aromaticity and pericyclic reactions. *Angew. Chem., Int. Ed.* **1971**, *10*, 761−776.

(4) Houk, K. N.; Li, Y.; Evanseck, J. D. Transition structures of hydrocarbon pericyclic reactions. *Angew. Chem., Int. Ed.* **1992**, *31*, 682−708.

(5) For the first report on stereospecific rotations in electrocyclic reactions, see: Woodward, R. B.; Hoffmann, R. Stereochemistry of electrocyclic reaction. *J. Am. Chem. Soc.* **1965**, *87*, 395−397.

(6) For Houk's reviews on pericyclic reactions, see: (a) Houk, K. N.; Gonzalez, J.; Li, Y. Pericyclic reaction transition states: passions and punctilios, 1935-1995. *Acc. Chem. Res.* **1995**, *28*, 81−90. (b) Guner, V.; Khuong, K. S.; Leach, A. G.; Lee, P. S.; Bartberger, M. D.; Houk, K. N. A standard set of pericyclic reactions of hydrocarbons for the benchmarking of computational methods: the performance of ab initio, density functional, CASSCF, CASPT2, and CBS-QB3 methods for the prediction of activation barriers, reaction energetics, and transition state geometries. *J. Phys. Chem. A* **2003**, *107*, 11445−11459. (c) Jamieson, C. S.; Ohashi, M.; Liu, F.; Tang, Y.; Houk, K. N. The expanding world of biosynthetic pericyclases: cooperation of experiment and theory for discovery. *Nat. Prod. Rep.* **2019**, *36*, 698−713. (d) McLeod, D.; Thøgersen, M. K.; Jessen, N. I.; Jørgensen, K. A.; Jamieson, C. S.; Xue, X.-S.; Houk, K. N.; Liu, F.; Hoffmann, R. Expanding the frontiers of higher-order cycloadditions. *Acc. Chem. Res.* **2019**, *52*, 3488−3501. (e) Houk, K. N.; Liu, F.; Yang, Z.; Seeman, J. I. Evolution of the Diels−Alder reaction mechanism since the 1930s: Woodward, Houk with Woodward, and the influence of computational chemistry on understanding cycloadditions. *Angew. Chem., Int. Ed.* **2021**, *60*, 12660−12681.

(7) (a) Schlegel, H. B. Exploring potential energy surfaces for chemical reactions: An overview of some practical methods. *J. Comput. Chem.* **2003**, *24*, 1514−1527. (b) Maeda, S.; Ohno, K. Global mapping of equilibrium and transition structures on potential energy surfaces by the scaled hypersphere search method: applications to ab initio surfaces of formaldehyde and propyne molecules. *J. Phys. Chem. A* **2005**, *109*, 5742−5753. (c) Maeda, S.; Morokuma, K. Communications: a systematic method for locating transition structures of A + B → X type reactions. *J. Chem. Phys.* **2010**, *132*, 241102. (d) Zimmerman, P. M. Automated discovery of chemically reasonable elementary reaction steps. *J. Comput. Chem.* **2013**, *34*, 1385−1392. (e) Rappoport, D.; Galvin, C. J.; Zubarev, D. Y.; Aspuru-Guzik, A. Complex chemical reaction networks from heuristics-aided quantum chemistry. *J. Chem. Theory Comput.* **2014**, *10*, 897−907. (f) Kim, Y.; Choi, S.; Kim, W. Y. Efficient basin-hopping sampling of reaction intermediates through molecular fragmentation and graph theory. *J. Chem. Theory Comput.* **2014**, *10*, 2419−2426. (g) Suleimanov, Y. V.; Green, W. H. Automated discovery of elementary chemical reaction steps using freezing string and Berny optimization methods. *J. Chem. Theory Comput.* **2015**, *11*, 4248−4259. (h) Zimmerman, P. M. Single-ended transition state finding with the growing string method. *J. Comput. Chem.* **2015**, *36*, 601−611. (i) Martínez-Núñez, E. An automated method to find transition states using chemical dynamics simulations. *J. Comput. Chem.* **2015**, *36*, 222−234. (j) Bergeler, M.; Simm, G. N.; Proppe, J.; Reiher, M. Heuristics-guided exploration of reaction mechanisms. *J. Chem. Theory Comput.* **2015**, *11*, 5712−5722. (k) Zhang, X.-J.; Liu, Z.-P. Reaction sampling and reactivity prediction using the stochastic surface walking method. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2757−2769. (l) Wang, L.-P.; McGibbon, R. T.; Pande, V. S.; Martinez, T. J. Automated discovery and refinement of reactive molecular dynamics pathways. *J. Chem. Theory Comput.* **2016**, *12*, 638−649. (m) Van de Vijver, R.; Zádor, J. KinBot: automated stationary point search on potential energy surfaces. *Comput. Phys. Commun.* **2020**, *248*, No. 106947.

(8) For reviews, see: (a) 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.* **2013**, *15*, 3683−3701. (b) Dewyer, A. L.; Argüelles, A. J.; Zimmerman, P. M. Methods for exploring reaction space in molecular systems. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1354. (c) Simm, G. N.; Vaucher, A. C.; Reiher, M. Exploration of reaction pathways and chemical transformation networks. *J. Phys. Chem. A* **2019**, *123*, 385−399.

(9) Grambow, C. A.; Jamal, A.; Li, Y.-P.; Green, W. H.; Zádor, J.; Suleimanov, Y. V. Unimolecular reaction pathways of a γ-ketohydroperoxide from combined application of automated reaction discovery methods. *J. Am. Chem. Soc.* **2018**, *140*, 1035−1048.

(10) For the AFIR method implemented in the GRRM program, see: Maeda, S.; Harabuchi, Y. Exploring paths of chemical transformations in molecular and periodic systems: An approach utilizing force. *WIREs Comput. Mol. Sci.* **2021**, *11*, No. e1538.

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

(12) (a) Mita, T.; Harabuchi, Y.; Maeda, S. Discovery of a synthesis method for a difluoroglycine derivative based on a path generated by quantum chemical calculations. *Chem. Sci.* **2020**, *11*, 7569−7577. (b) Hayashi, H.; Takano, H.; Katsuyama, H.; Harabuchi, Y.; Maeda, S.; Mita, T. Synthesis of difluoroglycine derivatives from amines, difluorocarbene, and CO_2: computational design, scope, and applications. *Chem. − Eur. J.* **2021**, *27*, 10040−10047. (c) Hayashi, H.; Katsuyama, H.; Takano, H.; Harabuchi, Y.; Maeda, S.; Mita, T. In silico reaction screening with difluorocarbene for *N*-difluoroalkylative dearomatization of pyridines. *Nat. Synth.* **2022**, *1*, 804−814.

(13) Sumiya, Y.; Harabuchi, Y.; Nagata, Y.; Maeda, S. Quantum chemical calculations to trace back reaction paths for the prediction of reactants. *JACS Au* **2022**, *2*, 1181−1188.

(14) (a) Nicolaou, K. C.; Petasis, N. A.; Zipkin, R. E.; Uenishi, J. The endiandric acid cascade. Electrocyclizations in organic synthesis. 1. Stepwise, stereocontrolled total synthesis of endiandric acids A and B. *J. Am. Chem. Soc.* **1982**, *104*, 5555−5557. (b) Nicolaou, K. C.; Petasis, N. A.; Uenishi, J.; Zipkin, R. E. The endiandric acid cascade. Electrocyclizations in organic synthesis. 2. Stepwise, stereocontrolled total synthesis of endiandric acids C-G. *J. Am. Chem. Soc.* **1982**, *104*, 5557−5558. (c) Nicolaou, K. C.; Zipkin, R. E.; Petasis, N. A. The endiandric acid cascade. Electrocyclizations in organic synthesis. 3. "Biomimetic" approach to endiandric acids A-G. Synthesis of precursors. *J. Am. Chem. Soc.* **1982**, *104*, 5558−5560. (d) Nicolaou, K. C.; Petasis, N. A.; Zipkin, R. E. The endiandric acid cascade. Electrocyclizations in organic synthesis. 4. Biomimetic approach to endiandric acids A-G. Total synthesis and thermal studies. *J. Am. Chem. Soc.* **1982**, *104*, 5560−5562.

(15) (a) Bandaranayake, W. M.; Banfield, J. E.; Black, D. S.; Fallon, G. D.; Gatehouse, B. M. Endiandric acid, a novel carboxylic acid from *Endiandra introrsa* (Lauraceae): X-ray structure determination. *J. Chem. Soc., Chem. Commun.* **1980**, 162−163. (b) Bandaranayake, W. M.; Banfield, J. E.; Black, D. S. Postulated electrocyclic reactions leading to endiandric acid and related natural products. *J. Chem. Soc., Chem. Commun.* **1980**, 902−903.

(16) Kukier, G. A.; Turlik, A.; Xue, X.-S.; Houk, K. N. Violations. How nature circumvents the Woodward−Hoffmann rules and promotes the forbidden conrotatory 4n + 2 electron electrocyclization of Prinzbach's vinylogous sesquifulvalene. *J. Am. Chem. Soc.* **2021**, *143*, 21694−21704.

(17) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revision C.01*; Gaussian, Inc.: Wallingford CT, 2016.

(18) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.

(19) 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* **2011**, *115*, 14556−14562.

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000

(20) (a) Sumiya, Y.; Maeda, S. A reaction path network for Wöhler's urea synthesis. *Chem. Lett.* **2019**, *48*, 47−50. (b) Sumiya, Y.; Maeda, S. Rate constant matrix contraction method for systematic analysis of reaction path networks. *Chem. Lett.* **2020**, *49*, 553−564.
(21) (a) Choi, C.; Elber, R. Reaction path study of helix formation in tetrapeptides: Effect of side chains. *J. Chem. Phys.* **1991**, *94*, 751−760. (b) Ayala, P. Y.; Schlegel, H. B. A combined method for determining reaction paths, minima, and transition state geometries. *J. Chem. Phys.* **1997**, *107*, 375−384.
(22) (a) Fukui, K. The path of chemical reactions - the IRC approach. *Acc. Chem. Res.* **1981**, *14*, 363−368. (b) Maeda, S.; Harabuchi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. Intrinsic reaction coordinate: Calculation, bifurcation, and automated search. *Int. J. Quantum Chem.* **2015**, *115*, 258−269.
(23) (a) Ellis, R. J.; Frey, H. M. The thermal unimolecular isomerisation of bicyclo[3,1,0] hex-2-ene and decomposition of cyclohexa-l,4-diene. *J. Chem. Soc. A* **1966**, 553−556. (b) Benson, S. W.; Shaw, R. Kinetics and mechanism of the pyrolysis of 1,3-cyclohexadiene. A thermal source of cyclohexadienyl radicals and hydrogen atoms. The addition of hydrogen atoms to benzene and toluene. *J. Am. Chem. Soc.* **1967**, *89*, 5351−5354.
(24) Mock, W. L. The stereoelectronic course of the diene-sulfur dioxide reaction. *J. Am. Chem. Soc.* **1966**, *88*, 2857−2858.
(25) Mock, W. L. Stereoelectronic course of the triene-sulfur dioxide reaction. *J. Am. Chem. Soc.* **1969**, *91*, 5682−5684.
(26) Coldham, I.; Burrell, A. J. M.; Guerrand, H. D. S.; Oram, N. Cascade cyclization, dipolar cycloaddition to bridged tricyclic amines related to the daphniphyllum alkaloids. *Org. Lett.* **2011**, *13*, 1267−1269.
(27) Kremsner, J. M.; Kappe, C. O. Silicon carbide passive heating elements in microwave-assisted organic synthesis. *J. Org. Chem.* **2006**, *71*, 4651−4658.
(28) Doering, W. E.; Roth, W. R. The overlap of two allyl radicals or a four-centered transition state in the Cope rearrangement. *Tetrahedron* **1962**, *18*, 67−74.
(29) Shea, K. J.; Phillips, R. B. Diastereomeric transition states. Relative energies of the chair and boat reaction pathways in the Cope rearrangement. *J. Am. Chem. Soc.* **1980**, *102*, 3156−3162.
(30) Skropeta, D.; Rickards, R. W. Domino pericyclic reactions of acyclic conjugated (*E,Z,E,E*)-tetraenes. *Tetrahedron Lett.* **2007**, *48*, 3281−3284.
(31) Yahiaoui, O.; Almass, A.; Fallon, T. Total synthesis of endiandric acid J and beilcyclone A from cyclooctatetraene. *Chem. Sci.* **2020**, *11*, 9421−9425.
(32) Prinzbach, H.; Babsch, H.; Hunkler, D. The "vinylogous sesquifulvalene" 14-electron-electrocyclisation reaction. *Tetrahedron Lett.* **1978**, *19*, 649−652.
(33) Monnat, F.; Vogel, P.; Sordo, J. Á. Hetero-Diels-Alder and cheletropic additions of sulfur dioxide to 1,2-dimethylidenecycloalkanes. Determination of thermochemical and kinetics parameters for reactions in solution and comparison with estimates from quantum calculations. *Helv. Chim. Acta* **2002**, *85*, 712−732.

https://doi.org/10.1021/jacs.2c09830
J. Am. Chem. Soc. 2022, 144, 22985−23000