Involvement of Lipocalin-like CghA in Decalin-Forming Stereoselective Intramolecular [4+2] Cycloaddition

ARTICLE *in* CHEMBIOCHEM · SEPTEMBER 2015

Impact Factor: 3.09 · DOI: 10.1002/cbic.201500386

ChemPubSoc Europe

DOI: 10.1002/cbic.201500386

CHEMBIOCHEM Communications

Involvement of Lipocalin-like CghA in Decalin-Forming Stereoselective Intramolecular [4+2] Cycloaddition

Michio Sato,[a] Fumitoshi Yagishita,[b] Takashi Mino,[b] Nahoko Uchiyama,[c] Ashay Patel,[d] Yit-Heng Chooi,[e] Yukihiro Goda,[c] Wei Xu,[e] Hiroshi Noguchi,[a] Tsuyoshi Yamamoto,[a] Kinya Hotta,[f] Kendall N. Houk,[d] Yi Tang,[e] and Kenji Watanabe*[a]

Understanding enzymatic Diels–Alder (DA) reactions that can form complex natural product scaffolds is of considerable interest. Sch 210972 **1**, a potential anti-HIV fungal natural product, contains a decalin core that is proposed to form through a DA reaction. We identified the gene cluster responsible for the biosynthesis of **1** and heterologously reconstituted the biosynthetic pathway in *Aspergillus nidulans* to characterize the enzymes involved. Most notably, deletion of *cghA* resulted in a loss of stereoselective decalin core formation, yielding both an *endo* (**1**) and a diastereomeric *exo* adduct of the proposed DA reaction. Complementation with *cghA* restored the sole formation of **1**. Density functional theory computation of the proposed DA reaction provided a plausible explanation of the observed pattern of product formation. Based on our study, we propose that lipocalin-like CghA is responsible for the stereoselective intramolecular [4+2] cycloaddition that forms the decalin core of **1**.

Although the Diels–Alder (DA) reaction, including catalytic variants of the reaction, has seen great use in chemical synthesis, biogenic and biocatalytic DA reactions are less well characterized. However, designed proteins[1] and nucleic acids[2] are capable of catalyzing the DA reaction, suggesting that biomolecules can act as competent catalysts for this cycloaddition. The DA reaction has been established as a key transformation in biosynthesis of a growing number of natural products.[3] Despite these proposals, the identification of enzymes, Diels–Alderases (DAases), that catalyze the cycloaddition has proven challenging. In fact, only a handful of DAases have been identified, based upon biochemical and structural studies, including lovastatin nonaketide synthase (LovB),[4] solanapyrone synthase,[5] SpnF (involved in the biosynthesis of spinosyn A),[6] and most recently, VstJ from the versipelostatin biosynthetic pathway (responsible for the formation of spirotetronate framework).[7] Of these examples, SpnF and VstJ are the only enzymes for which specific rate accelerations of [4+2] cycloaddition reactions have been verified experimentally. Given the limited understanding of enzymatic DA reactions and the potential synthetic utility of (re)engineered DAases, additional studies establishing the existence and catalytic modes of natural DAases are warranted (Scheme 1).

[a] *Dr. M. Sato, Prof. Dr. H. Noguchi, T. Yamamoto, Prof. Dr. K. Watanabe*
Department of Pharmaceutical Sciences, University of Shizuoka
52–1 Yada, Suruga-ku, Shizuoka 422-8526 (Japan)
E-mail: kenji55@u-shizuoka-ken.ac.jp

[b] *Dr. F. Yagishita, Prof. Dr. T. Mino*
Department of Applied Chemistry and Biotechnology
Graduate School of Engineering, Chiba University
1–33 Yayoi-cho, Inage-ku, Chiba 263-8522 (Japan)

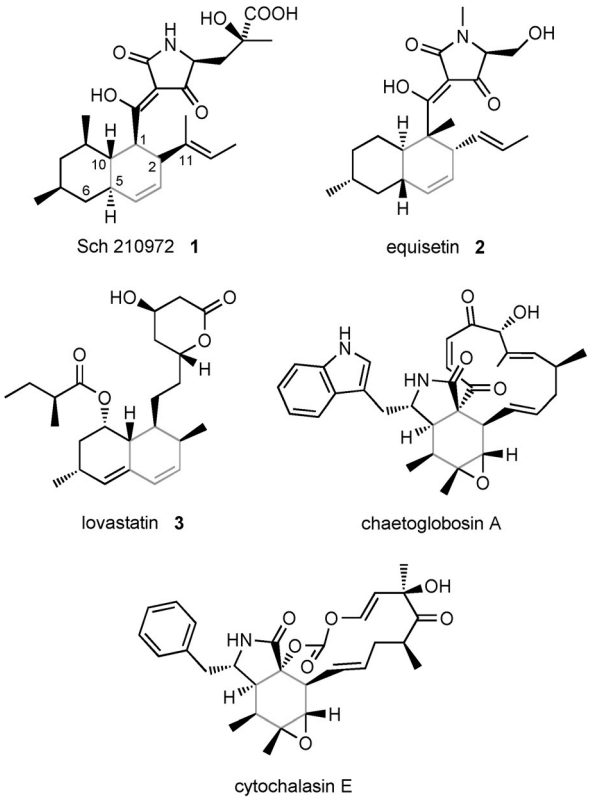
[c] *Dr. N. Uchiyama, Dr. Y. Goda*
National Institute of Health Sciences
1–18–1 Kamiyoga, Setagaya-ku, Tokyo 158-8501 (Japan)

[d] *Dr. A. Patel, Prof. Dr. K. N. Houk*
Department of Chemistry and Biochemistry, University of California
607 Charles E. Young Drive East, Box 951569, Los Angeles, CA 90095 (USA)

[e] *Dr. Y.-H. Chooi, Dr. W. Xu, Prof. Dr. Y. Tang*
Department of Chemical and Biomolecular Engineering
and Department of Chemistry and Biochemistry, University of California
420 Westwood Plaza, Los Angeles, CA 90095 (USA)

[f] *Prof. Dr. K. Hotta*
School of Biosciences, The University of Nottingham Malaysia Campus
Jalan Broga, Selangor 43500 (Malaysia)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201500386.



Scheme 1. Natural products putatively biosynthesized by Diels–Alder reactions. Resulting cyclohexyl groups are shown in gray.

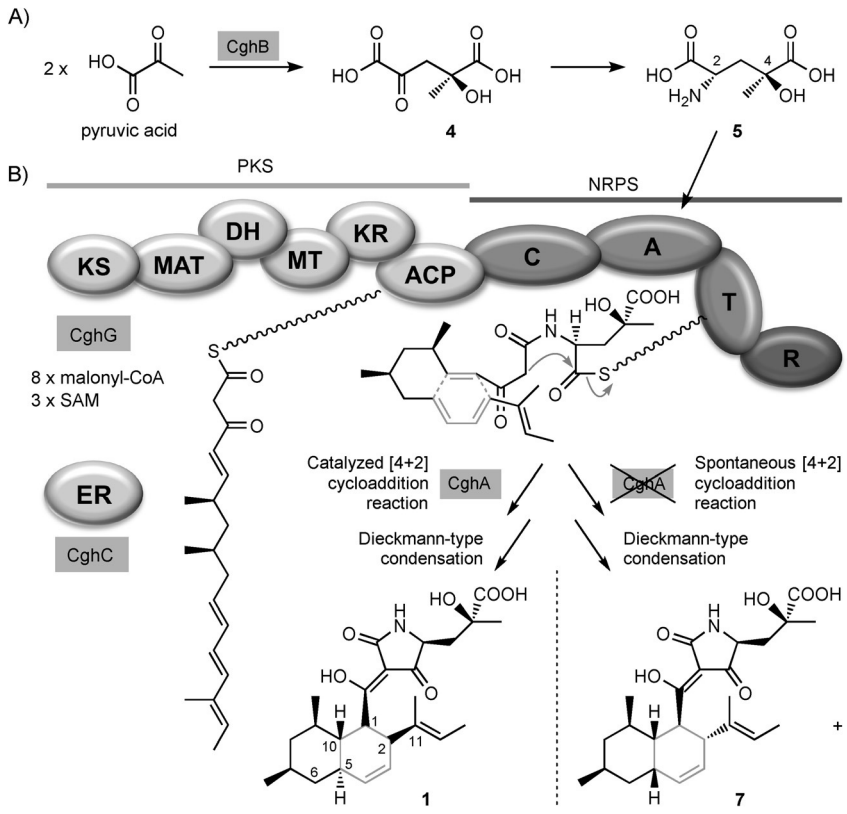
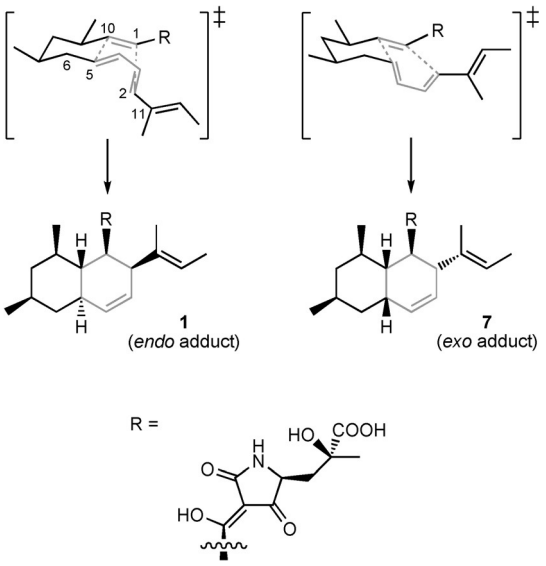


Figure 1. Proposed biosynthetic pathway of **1**. A) Formation of an unusual amino acid, γ-hydroxymethyl-L-glutamic acid **5**, from two molecules of pyruvic acid dimerized by an aldolase, CghB. B) CghG-catalyzed synthesis of the linear substrate and a tetramic acid moiety, and the two final products that can be formed by the proposed Diels–Alder reaction. Abbreviations: KS: ketosynthase, MAT: malonyl-CoA acyltransferase, DH: dehydratase, MT: methyltransferase, KR: ketoreductase, ACP: acyl carrier protein, C: condensation, A: adenylation, T: thiolation, R: reductase, ER: enoyl reductase, SAM: *S*-adenosyl-L-methionine.

We have investigated the biosynthetic pathway responsible for the formation of the tetramic acid-containing metabolite Sch 210972 **1** (Figure 1) from *Chaetomium globosum*. This compound exhibits unique inhibitory activity against CCR-5, a cell surface receptor that allows the entry of HIV-1 into cells, and is an attractive lead compound for novel anti-HIV therapeutics.[8] Previous studies have shown that the tetramic acid moieties in related compounds, such as equisetin **2**[9] (Scheme 1) and cyclopiazonic acid,[10] are formed by a Dieckmann-type condensation that is catalyzed by reductive domains located at the C termini of polyketide synthase–nonribosomal peptide synthetase hybrid megaenzymes (PKS–NRPSs).[11] Thus, we hypothesized that the framework of **1** was also biosynthesized by a hybrid PKS–NRPS. The structural similarity of **1**, **2**, and lovastatin **3** suggested that the decalin core of **1** might be formed by an intramolecular DA reaction, as proposed for the biocatalytic formation of the decalin cores of **2** and **3** (Figure 1 B).[12] To gain further insight into the mechanism of biosynthesis of **1** and to determine if a DA reaction is indeed involved in the biosynthesis of **1**, we identified the *C. globosum* gene cluster responsible for the biosynthesis of **1** for a detailed investigation. (Scheme 2)

A BLASTP[13] search of the genome of *C. globosum* identified three recognizable hybrid PKS–NRPS genes: CHGG_01239, CHGG_05286, and CHGG_02374–CHGG_02378. We have recently reported that CHGG_01239 codes for the PKS–NRPS CheA responsible for the formation of chaetoglobosin-type of natural products[14] (Figure S13), and deletion of CHGG_05286 did not affect the production of **1**. Therefore, the reassigned ORF CHGG_02374-CHGG_02378, which we have named *cghG*, was predicted and later confirmed to code for the PKS–NRPS responsible for the biosynthesis of **1** (see Table S2, Figures S1–S9; S14, i vs. viii in the Supporting Information). The chemical structure of **1** was characterized by using high-resolution electrospray ionization (HR-ESI) LCMS, ^{1}H NMR, and ^{13}C NMR (Table S3 and Figure S16, S17), and the absolute configuration of **1** was determined crystallographically (Figure S18 and Tables S5–S12). Additionally, this class of PKS–NRPS is often accompanied by a stand-alone enoyl reductase (ER) that works in *trans* with the PKS–NRPS for reduction of polyketide backbone double bonds that is required for the formation of the final products.[11,15] Based on homology to CheB, the ER involved in chaetoglobosin biosynthesis,[16] CghC was predict-



Scheme 2. The *endo* and *exo* transition state of the [4+2] cycloaddition, leading to the formation of **1** and **7**, respectively.

ed to be the stand-alone ER for the biosynthesis of **1** (Table S2), and its disruption abolished the production of **1** (Figure S14, i vs. iv). Lastly, disruption of *cghD* also showed a complete loss of production of **1** (Figure S14, i vs. v), presumably because CghD is the GAL4-type transcription factor that activates expression of the *cgh* genes.

One unique structural feature of **1** is the tetramic acid motif proposed to be derived from a non-proteinogenic amino acid, γ-hydroxymethyl-L-glutamic acid **5**, by a Dieckmann-type condensation catalyzed by the reductase domain of CghG (Figure 1 B), as in the case of the biosynthesis of **2**.[17] Because the *cgh* gene cluster possessed an aldolase-coding *cghB*, we proposed that CghB catalyzes the aldol condensation of two molecules of pyruvic acid to yield the intermediate **4**, which can then be stereoselectively transaminated to form **5** (Figure 1 A). To test this hypothesis, we prepared **5** and its diastereomers chemoenzymatically[18] and used them in feeding experiments with an engineered *C. globosum* strain, CGKW14[19] (Supporting Information). After confirming the loss of production of **1** by *ΔcghB*/CGKW14 (Figure S14, i vs. iii), we fed **5** to this strain to observe the recovery of the production of **1** (Figure 2). Amination of **4** is likely performed by a transaminase in the cell. On the other hand, feeding of a diastereomer of **5**, (2*S*,4*R*)-4-hydroxy-4-methylglutamic acid **6** (Supporting Methods), failed to restore the production of **1** in *ΔcghB*/CGKW14 (Figure S15). These results suggest that CghB is responsible for supplying **5** to CghG to complete the biosynthesis of **1** and also revealed the stereospecificity of the adenylation domain of the NRPS segment of CghG.

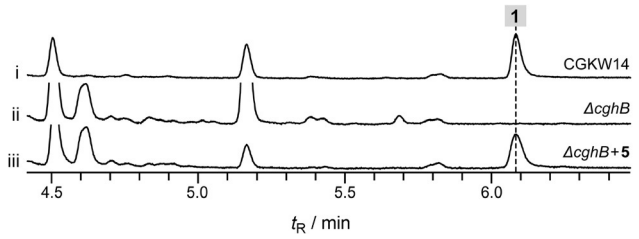


Figure 2. Confirmation of the activity of CghB by bioconversion of synthetic substrate **5** to **1** in *C. globosum*. HPLC traces of extracts of the culture of i) wild-type control CGKW14, ii) negative control of *ΔcghB*/CGKW14 that was not fed **5** and iii) *ΔcghB*/CGKW14 that was fed **5**.

The biosynthetic gene clusters for all of the compounds except **3** in Scheme 1 contain a gene encoding hypothetical proteins that are homologous to one another (Figure S13). For **1**, this gene corresponds to *cghA* (Table S2). Although CcsF, the homologue of CghA in the cytochalasin E biosynthetic pathway, was suggested to play a possible role in the formation of the isoindolone core by catalyzing a DA reaction,[15b] none of those *cghA*-like genes has been examined experimentally until now. When *cghA* was disrupted, **1** was still formed, albeit in a reduced quantity (Figure S14, i vs. ii), but also resulted in the formation of a second compound, **7** (Figure 3, ii). The wild-type CGKW14 strain did not form **7** (Figure 3, i). Detailed characterizations by HR-ESI-LCMS, ^{1}H NMR, and ^{13}C NMR (Table S4 and Figures S19–S24), as well as X-ray crystallography (Figure S18), determined **7** to be a diastereomer of **1** where **1** and **7** correspond to the *endo* and *exo* adduct of the proposed decalin-forming [4+2] cycloaddition reaction, respectively. The transition states leading to these cycloadducts are shown in Scheme 2.

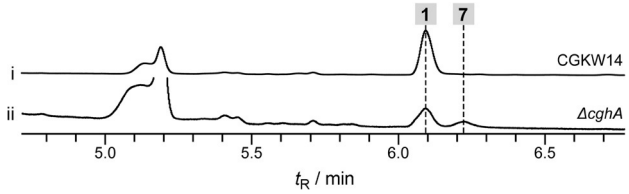


Figure 3. HPLC analysis of the metabolites from *ΔcghA* strain of *C. globosum*. HPLC traces of extracts of the culture of i) wild-type control CGKW14 and ii) deletion mutant *ΔcghA*/CGKW14. The yields of **1** and **7** from the *ΔcghA*/CGKW14 culture were 2 and 1 mg L^{-1}, respectively, whereas 60 mg L^{-1} of **1** was isolated from the wild-type reference CGKW14 culture.

To confirm the involvement of CghA in dictating the course of the stereoselective intramolecular [4+2] cycloaddition reaction, we conducted a complementation experiment in which a functional copy of the *cghA* gene was reintroduced to the *ΔcghA*/CGKW14 strain. We chose to integrate *cghA* into the chaetoviridin biosynthetic PKS gene *cazM*[20] so that the integration event would be indicated by the loss of formation of chaetoviridin A (**8**) and B (**9**).[19] A control experiment verified that disruption of *cazM* would not affect the production of **1** (Figure 4, i vs. ii). Comparison of the extract from the *cghA* knockout strain and the *cghA*-complemented strain showed that the biosynthesis of **7** was no longer observed and exclusive production of **1** was restored with *cghA* (Figure 4, iii vs. iv), indicating that CghA is directly responsible for ensuring the stereoselective formation of **1** and suppressing the nonenzymatic formation of the alternative stereoisomer, **7**. Furthermore, the Sch 210972 biosynthetic pathway was reconstituted

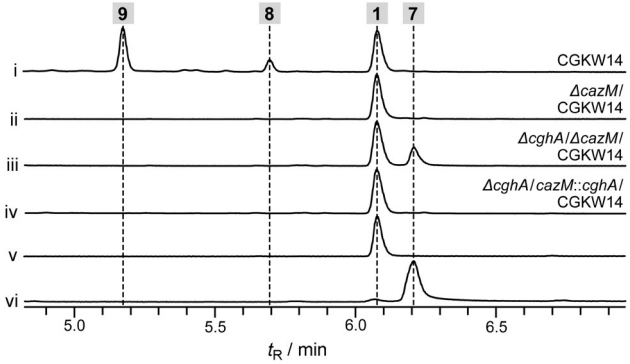


Figure 4. Confirmation of the role of CghA in the selective formation of **1** by complementing the *cghA* disruption mutant with *cghA* in *C. globosum*. HPLC traces of extracts of the culture of i) wild-type control CGKW14, ii) *ΔcazM*/CGKW14 as a control showing the loss of production of **8** and **9**, iii) *ΔcghA*/*ΔcazM*/CGKW14, and iv) *ΔcghA*/*cazM*::*cghA*/CGKW14. Authentic reference of v) **1** and vi) **7**.

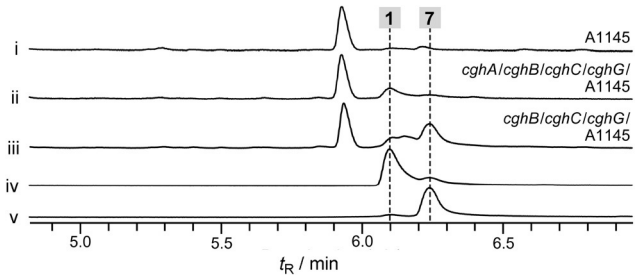


Figure 5. Reconstitution of the biosynthesis of **1** in *Aspergillus nidulans* A1145 and loss of stereoselective product formation upon elimination of *cghA*. HPLC traces of metabolic extracts from the culture of i) wild-type *A. nidulans* A1145 control, ii) *A. nidulans* harboring *cghA*, *cghB*, *cghC*, and *cghG*, and iii) *A. nidulans* harboring *cghB*, *cghC*, and *cghG*. Authentic reference of iv) **1** and v) **7**.

in a heterologous host to eliminate any interfering effects in *C. globosum* that could obscure our findings (Supporting Information). As expected, introduction of *cghA*, *cghB*, *cghC*, and *cghG* to *Aspergillus nidulans* A1145 resulted in the formation of **1** (Figure 5 i vs. ii), whereas exclusion of *cghA* led to the formation of **1** and **7** (Figure 5 iii). These results clearly indicate that CghA selectively promotes a single stereochemical pathway among the decalin-forming intramolecular [4+2] cycloaddition reaction pathways to yield **1** exclusively.

Lastly, to quantitate the stereoselectivity of the cycloaddition, we computed the DA reaction of model substrates **10** and **11** shown in Figure 6 A and Figure S25 by using the M06-2X density functional and 6–31+G(d,p) basis with the SMD implicit solvation model. Computations of the reaction with both substrates led to similar conclusions. Therefore, the discussion here will focus on the computations of the DA reaction of truncated substrate **10**. Details regarding density functional theory computations of the cycloaddition of **10** and **11** and our computational methods are presented in the Supporting Methods. The activation free energies and the transition structures for the formation of the four stereoisomeric cycloadduct of **10** are shown in Figure 6 B. The barriers for the formation of these stereoadducts ranged from 18 to 27 kcal mol^{-1}. TS10$_{endo1}$ and TS10$_{exo1}$, the lowest energy transition structures, led to cycloadducts **12a** and **12b**, which are analogous in terms of stereochemistry to the adducts **1** and **7** formed by the *ΔcghA* strain. The dominant adduct formed in the *ΔcghA* strain was **1**, with **7** being a less prevalent product (Figure 3 ii), and our computations reproduced this selectivity trend correctly. The transition state for the formation of **12a** (TS10$_{endo1}$) is 1.3 kcal mol^{-1} lower in energy than the transition state leading to **12b** (TS10$_{exo1}$), suggesting approximately a 10:1 preference for the formation of **12a**. TS10$_{exo1}$ is destabilized by a 1,3-diaxial interaction involving the methylene group of the diene. Similar, but more severe 1,3-diaxial interactions destabilize the other *endo* and *exo* transition states (TS10$_{exo2}$ and TS10$_{endo2}$), thus causing them to be more than 4 kcal mol^{-1} higher in energy than TS10$_{endo1}$.

Sequence analysis of CghA and its homologues by FFAS03[21] identified a calycin-like β-barrel protein, NE1406,[22] as the closest structural homologue of CghA. Another family of calycin-

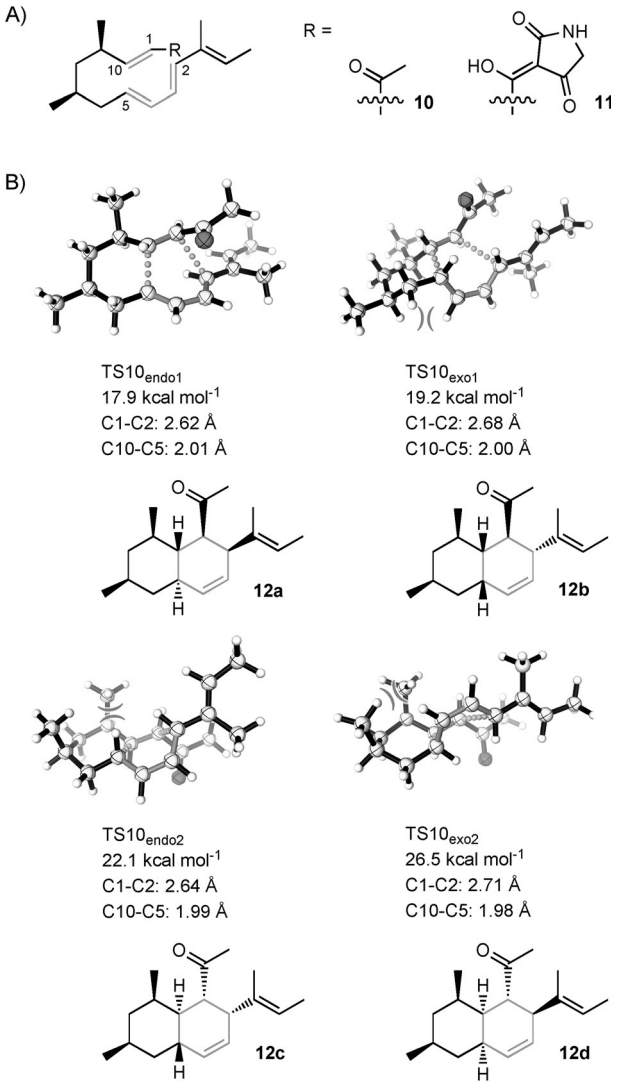


Figure 6. Computational studies of the proposed Diels–Alder reaction involved in the biosynthesis of Sch 210972. A) Truncated substrates examined by using density functional theory. B) Computed transition structures of the intramolecular Diels–Alder reaction of **10**. The activation free energy (bold) and the forming bond lengths are given. Unfavorable 1,3-diaxial interactions are indicated by opposing arcs.

like β-barrel proteins, called dirigent proteins, is known to sequester and orient phenoxy radicals to direct stereospecific formation of dilignols, precursors to lignin and lignans.[23] In a similar fashion, CghA-like proteins might also bind specific conformations of their straight-chain polyketide substrates to promote the [4+2] cycloaddition reaction and to control the stereoselectivity of the reaction. Computations indicate that catalysis by substrate preorganization could accelerate the cycloaddition by as much a 1000-fold at 30 °C, as the reactive conformer for the formation of **12a** is 4 kcal mol^{-1} higher in energy than the global minimum conformer. Hydrogen bonding of the enone carbonyl in the active site of CghA could also activate the precursor and further accelerate the cycloaddition. Although detailed in vitro characterization of the biosynthesis of **3**[4] and the recent identification of SpnF[6] and VstJ[7] strongly indicate the existence of natural enzymes that catalyze DA

reactions, little remains known about how exactly these enzymes catalyze asymmetric DA reactions. Efforts to further study the CghA-catalyzed intramolecular [4+2] cycloaddition reaction are currently underway to understand the possible involvement of DA-catalyzing enzymes in natural product biosynthesis.

Acknowledgements

We acknowledge financial support from Japan Society for the Promotion of Science Program for Advancing Strategic International Networks to Accelerate the Circulation of Talented Researchers (No. G2604), the Amano Enzyme Foundation, and the Nagase Science and Technology Foundation of Japan (K.W.). K.N.H acknowledges financial support from the National Science Foundation (NSF;(CHE-1059084) and the National Institutes of Health (NIH; 1R01M097200). A.P. thanks the Chemistry Biology Training Grant Program by the UCLA (T32 GM 008496) for their support and the University of California, Los Angeles for funding. We thank the UCLA Hoffman2 cluster and the NSF Extreme Science and Engineering Discovery Environment (TG-CHE040013N) for computing time on the Gordon and Stampede cluster. Y.T. is supported by the NIH (1DP1M106413).

Keywords: cycloaddition · decalin · density functional calculations · fungal metabolite · tetramic acid

[1] a) Y. Xu, N. Yamamoto, K. D. Janda, *Bioorg. Med. Chem.* **2004**, *12*, 5247 – 5268; b) N. Preiswerk, T. Beck, J. D. Schulz, P. Milovník, C. Mayer, J. B. Siegel, D. Baker, D. Hilvert, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8013 – 8018.

[2] A. Jäschke, B. Seelig, *Curr. Opin. Chem. Biol.* **2000**, *4*, 257 – 262.

[3] H. Oikawa in *Comprehensive Natural Products II, Vol. VIII*, (Eds.: H.-W. Liu, L. Mander), Elsevier, Oxford, **2010**, pp. 277 – 314.

[4] K. Auclair, A. Sutherland, J. Kennedy, D. J. Witter, J. P. Van den Heever, C. R. Hutchinson, J. C. Vederas, *J. Am. Chem. Soc.* **2000**, *122*, 11519 – 11520.

[5] K. Kasahara, T. Miyamoto, T. Fujimoto, H. Oguri, T. Tokiwano, H. Oikawa, Y. Ebizuka, I. Fujii, *ChemBioChem* **2010**, *11*, 1245 – 1252.

[6] H. J. Kim, M. W. Ruszczycky, S. H. Choi, Y. N. Liu, H. W. Liu, *Nature* **2011**, *473*, 109 – 112.

[7] T. Hashimoto, J. Hashimoto, K. Teruya, T. Hirano, K. Shin-Ya, H. Ikeda, H. W. Liu, M. Nishiyama, T. Kuzuyama, *J. Am. Chem. Soc.* **2015**, *137*, 572 – 575.

[8] S. W. Yang, R. Mierzwa, J. Terracciano, M. Patel, V. Gullo, N. Wagner, B. Baroudy, M. Puar, T. M. Chan, A. T. McPhail, M. Chu, *J. Nat. Prod.* **2006**, *69*, 1025 – 1028.

[9] T. B. Kakule, D. Sardar, Z. Lin, E. W. Schmidt, *ACS Chem. Biol.* **2013**, *8*, 1549 – 1557.

[10] X. Liu, C. T. Walsh, *Biochemistry* **2009**, *48*, 8746 – 8757.

[11] D. Boettger, C. Hertweck, *ChemBioChem* **2013**, *14*, 28 – 42.

[12] C. D. Campbell, J. C. Vederas, *Biopolymers* **2010**, *93*, 755 – 763.

[13] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, T. L. Madden, *Nucleic Acids Res.* **2008**, *36*, W5 – 9.

[14] K. Ishiuchi, T. Nakazawa, F. Yagishita, T. Mino, H. Noguchi, K. Hotta, K. Watanabe, *J. Am. Chem. Soc.* **2013**, *135*, 7371 – 7377.

[15] a) S. M. Ma, J. W. Li, J. W. Choi, H. Zhou, K. K. Lee, V. A. Moorthie, X. Xie, J. T. Kealey, N. A. Da Silva, J. C. Vederas, Y. Tang, *Science* **2009**, *326*, 589 – 592; b) K. Qiao, Y. H. Chooi, Y. Tang, *Metab. Eng.* **2011**, *13*, 723 – 732.

[16] J. Schümann, C. Hertweck, *J. Am. Chem. Soc.* **2007**, *129*, 9564 – 9565.

[17] J. W. Sims, E. W. Schmidt, *J. Am. Chem. Soc.* **2008**, *130*, 11149 – 11155.

[18] V. Hélaine, J. Rossi, T. Gefflaut, S. Alaux, J. Bolte, *Adv. Synth. Catal.* **2001**, *343*, 692 – 697.

[19] T. Nakazawa, K. Ishiuchi, M. Sato, Y. Tsunematsu, S. Sugimoto, Y. Gotanda, H. Noguchi, K. Hotta, K. Watanabe, *J. Am. Chem. Soc.* **2013**, *135*, 13446 – 13455.

[20] J. M. Winter, M. Sato, S. Sugimoto, G. Chiou, N. K. Garg, Y. Tang, K. Watanabe, *J. Am. Chem. Soc.* **2012**, *134*, 17900 – 17903.

[21] L. Jaroszewski, Z. Li, X. H. Cai, C. Weber, A. Godzik, *Nucleic Acids Res.* **2011**, *39*, W38 – 44.

[22] H. J. Chiu, C. Bakolitsa, A. Skerra, A. Lomize, D. Carlton, M. D. Miller, S. S. Krishna, P. Abdubek, T. Astakhova, H. L. Axelrod, T. Clayton, M. C. Deller, L. Duan, J. Feuerhelm, J. C. Grant, S. K. Grzechnik, G. W. Han, L. Jaroszewski, K. K. Jin, H. E. Klock, M. W. Knuth, P. Kozbial, A. Kumar, D. Marciano, D. McMullan, A. T. Morse, E. Nigoghossian, L. Okach, J. Paulsen, R. Reyes, C. L. Rife, H. van den Bedem, D. Weekes, Q. Xu, K. O. Hodgson, J. Wooley, M. A. Elsliger, A. M. Deacon, A. Godzik, S. A. Lesley, I. A. Wilson, *Acta Crystallogr. Sect. F* **2010**, *66*, 1153 – 1159.

[23] a) L. B. Davin, H. B. Wang, A. L. Crowell, D. L. Bedgar, D. M. Martin, S. Sarkanen, N. G. Lewis, *Science* **1997**, *275*, 362 – 366; b) B. Pickel, J. Pfannstiel, A. Steudle, A. Lehmann, U. Gerken, J. Pleiss, A. Schaller, *FEBS J.* **2012**, *279*, 1980 – 1993.

Manuscript received: August 26, 2015

Accepted article published: September 11, 2015

Final article published: October 2, 2015