Computationally Augmented Retrosynthesis: Total Synthesis of Paspaline A and Emindole PB

Daria E. Kim, Joshua E. Zweig and Timothy R. Newhouse*

Department of Chemistry
Yale University
225 Prospect Street
New Haven, Connecticut 06520-8107

Abstract: The diverse molecular architectures of terpene natural products are assembled by exquisite enzyme-catalyzed reactions. Successful recapitulation of these transformations using chemical synthesis is hard to predict from first principles and therefore challenging to execute. A means of evaluating the feasibility of such chemical reactions would greatly enable the development of concise syntheses of complex small molecules. Herein, we report the computational analysis of the energetic favorability of a key bio-inspired transformation, which we use to inform our synthetic strategy. This approach was applied to synthesize two constituents of the historically challenging indole diterpenoid class, resulting in a concise route to (–)-paspaline A in 9 steps from commercially available materials and the first pathway to and structural confirmation of emindole PB in 13 steps. This work highlights how traditional retrosynthetic design can be augmented with quantum chemical calculations to reveal energetically feasible synthetic disconnections, minimizing time-consuming and expensive empirical evaluation.

Text: Retrosynthetic analysis is the most widely utilized theoretical approach to designing syntheses of small molecules. It begins with the target structure and extrapolates backwards, step-by-step, until commercially available starting materials are reached[1]. Often, the most simplifying and valuable retrosynthetic disconnections (including some biomimetic processes) lack sufficient literature precedent to evaluate their feasibility without chemical experimentation. Modern computational tools offer powerful new ways to augment traditional retrosynthetic analysis. One of these approaches is the development of algorithms to expedite the identification of strategic bond disconnections[2–4]. An underutilized, yet complementary, strategy is to employ quantum chemical calculations to rigorously predict the feasibility of the most challenging and least precedented aspects of a synthetic plan[5,6]. Herein, we report the application of predictive density functional theory (DFT) calculations to the concise synthesis of two biosynthetically related, but structurally distinct indole diterpenoids, paspaline A (**1**) and emindole PB (**2**) (Fig. 1).

These molecules belong to a class of complex indole diterpenoid natural products that have been the targets of synthetic investigations for several decades[7–11]. The hexacyclic architecture of paspaline A (**1**), a flagship member of the class, includes a pyran ring as well as an indole-fused cyclopentane adjacent to two vicinal quaternary centers[12]. Emindole PB (**2**), a natural product with multiple reported structural assignments[13,14], has not been the subject of any reported synthetic efforts. Our interest in these compounds was motivated by the biological activity of paspaline A and related natural products, which includes anti-proliferative and anti-metastatic activity in the MDA-MB-231 mammalian breast cancer cell line through modulation of the Wnt/β–catenin signaling pathway[15]. The Wnt signaling pathway is a target of interest in solid tumor cancers with high rates of metastasis, such as breast cancer, due to its role in critical cellular processes like proliferation, differentiation and migration[16].

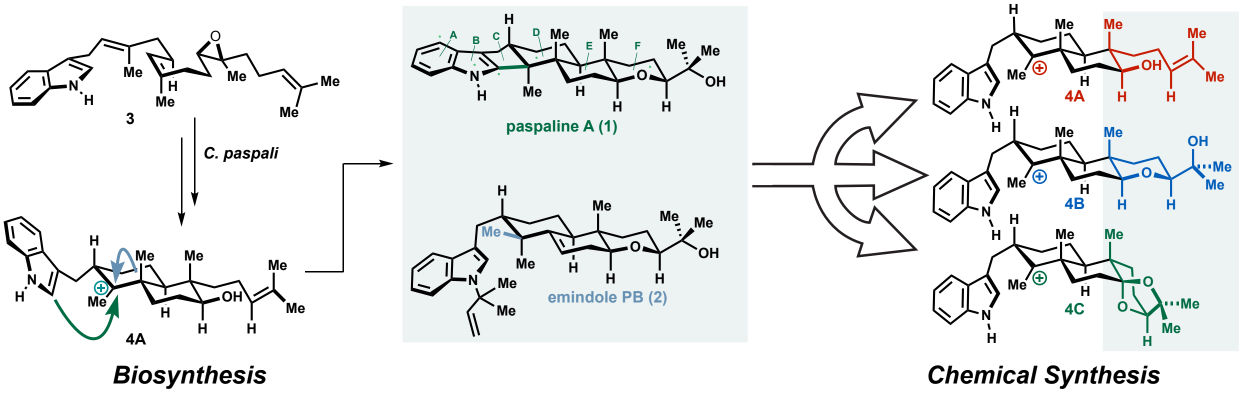
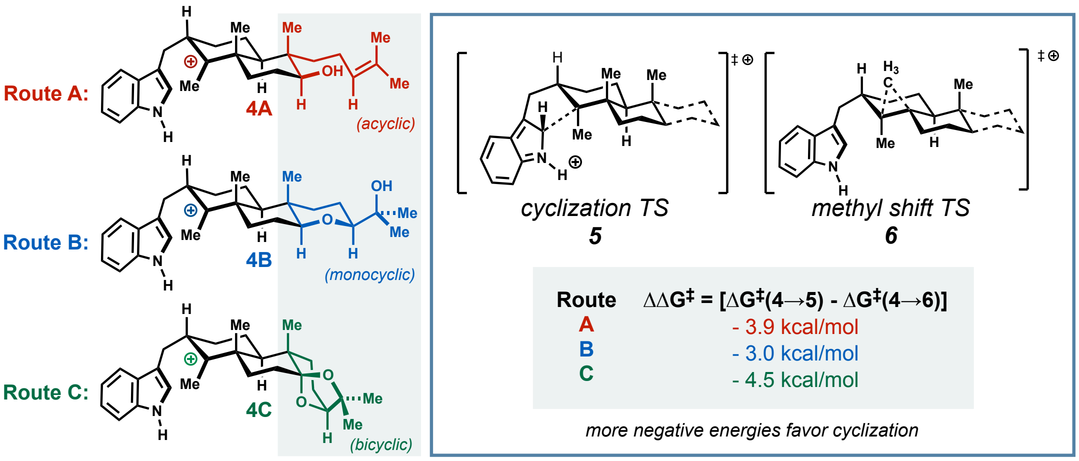


Figure 1. Computationally guided retrosynthesis. Biosynthesis of **1** and **2** via **3** and **4** suggests several potentially viable cyclization precursors for chemical synthesis (**4A–4C**).

Previous synthetic strategies towards these indole diterpenoids have relied on a late-stage indole ring synthesis by annulation onto an existing five-membered ring motif[7–9]. Inspired by the polyene cascade that assembles the indole diterpenoid skeleton in the putative biosynthesis, we planned to directly incorporate an indole motif and leverage its innate nucleophilicity to expedite synthesis of the C-ring (Fig. 1)[17,18]. The cascade proceeds through a notable biosynthetic fork in which the operative carbocation **4** is thought to give rise to three distinct subclasses of natural products via Friedel-Crafts cyclization to paspaline A (**1**), methyl migration to emindole PB (**2**), or elimination to give the anthcolorin class of natural products[19]. Synthetic recapitulation of this step would allow access to several structurally distinct members of the class in a single transformation.

In our retrosynthesis, we planned to modularly introduce the F-ring precursor and indole fragment by two alkylations of a Wieland-Miescher-ketone derivative, along with a strategic stereoselective alkene reduction. Within the framework of this general retrosynthetic plan, there is flexibility with respect to structural variation of the F-ring. We identified a set of substructures that would lead to step efficient routes, yet were structurally distinct enough to elicit measurable energetic differences with respect to the fate of the tertiary carbocation: the acyclic biosynthetic intermediate **4A**, a monocyclic tetrahydropyran **4B**, and a bicyclic ketal **4C** (Fig. 2). Despite these obvious structural differences, the distal nature of the modifications made it challenging to predict from first principles which of these substrates would optimally promote access to both the cyclization and 1,2-methyl shift. Therefore, we sought to use DFT calculations to refine our selection of the key carbocationic rearrangement precursor as a function of F-ring structure by comparing the energetic barriers for the two possible diastereomeric cyclizations and the methyl shift.

In considering which mechanistic pathway to model, we reasoned that the 1,2-methyl shift could occur via a concerted[20] or stepwise pathway, however the cyclization to form vicinal quaternary centers is presumably limited to the carbocationic pathway. Thus, the stepwise pathways were modeled by DFT to assess which of these three aforementioned substrates would be most optimal for cyclization rather than methyl shift. When the relative energies ($\Delta\Delta G^{\ddagger}$) between the gas phase transition states leading to cyclization and methyl-shift for each carbocationic substrate were compared at the mPW1PW91/6-31+G(d,p)//B3LYP/6-31G* level of theory[21], it was found that **4C** possessed the largest energetic bias (4.5 kcal/mol) towards cyclization via this stepwise pathway. As it is not feasible to calculate the potential range of exact experimental conditions, the calculations were limited to a simplified system (i.e. the gas phase carbocation) to provide a relative ranking of substrate viability rather than an exact distribution of products.



Route	$\Delta\Delta G^‡ = [\Delta G^‡(4 \rightarrow 5) - \Delta G^‡(4 \rightarrow 6)]$
A	- 3.9 kcal/mol
B	- 3.0 kcal/mol
C	- 4.5 kcal/mol

more negative energies favor cyclization

Figure 2. Computational evaluation of three variations on key structural reorganization. Three cyclization precursors with varied F-ring structures were investigated using DFT. These calculations predicted that **4C** would possess the largest substrate bias ($\Delta\Delta G^‡$) favoring C-ring cyclization over carbocation rearrangement. (Calculations at the mPW1PW91/6-31+G(d,p)//B3LYP/6-31G* level of theory.)

An additional challenge in effecting the desired indole cyclization reaction to form paspaline is that this ring formation would need to result in a *trans*-fused 5-membered ring, despite the fact that cyclization reactions that form fused 5-membered rings generally lead to the formation of *cis*-fused ring systems. This stereochemical outcome has even been observed in previous synthetic studies on closely related indoloterpenoid scaffolds which form *cis*-fused systems[22,23]. Fortuitously, DFT analysis of all four diastereomeric cyclization intermediates of precursor **4C** predicted that our desired stereochemical outcome (*trans*-C/D-fusion) was the more energetically favorable pathway by at least 1.3 kcal/mol (see Supporting Information Figure S2).

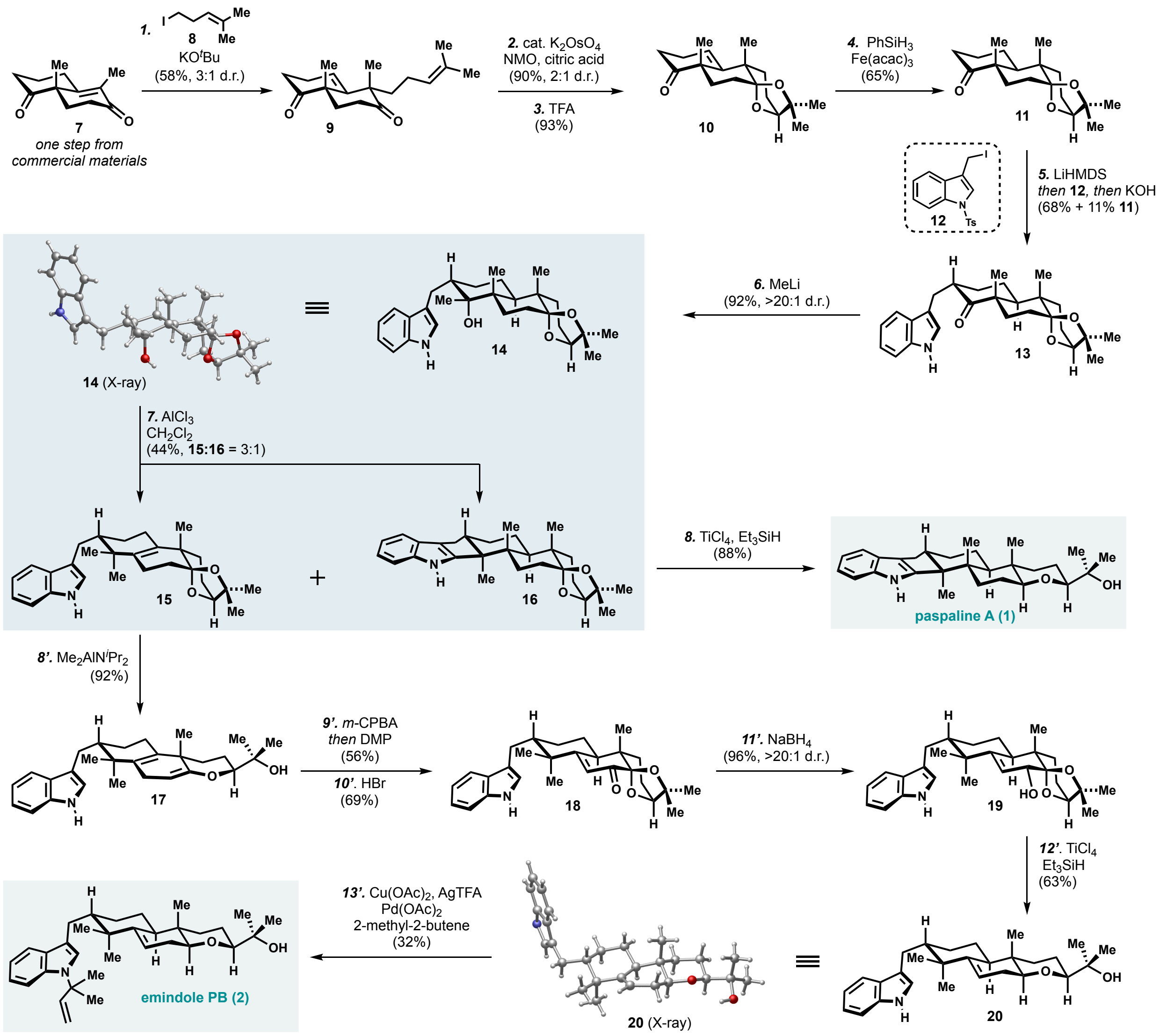

1. 8 KO^{t}Bu (58%, 3:1 d.r.)
7
one step from commercial materials
9
2. cat. K_2OsO_4 NMO, citric acid (90%, 2:1 d.r.)
3. TFA (93%)
10
4. $PhSiH_3$ $Fe(acac)_3$ (65%)
11
12
5. LiHMDS then 12, then KOH (68% + 11% 11)
13
6. MeLi (92%, >20:1 d.r.)
14
14 (X-ray)
7. $AlCl_3$ CH_2Cl_2 (44%, 15:16 = 3:1)
15
16
8. $TiCl_4$, Et_3SiH (88%)
paspaline A (1)
8'. $Me_2AlN^iPr_2$ (92%)
17
9'. m-CPBA then DMP (56%)
10'. HBr (69%)
18
11'. $NaBH_4$ (96%, >20:1 d.r.)
19
12'. $TiCl_4$ Et_3SiH (63%)
20
20 (X-ray)
13'. $Cu(OAc)_2$, AgTFA $Pd(OAc)_2$ 2-methyl-2-butene (32%)
emindole PB (2)

Figure 3. Total synthesis of (–)-paspaline A and (–)-emindole PB. Reagents and conditions as follows: 1. KO^{t}Bu (1.2 equiv), THF, 60 °C; then **8** (1.3 equiv), THF, 0 °C to 23 °C, 58% yield, 3:1 d.r. 2. K_2OsO_4 (2.5 mol%), citric acid (10 mol%), 4-methylmorpholine *N*-oxide (3.0 equiv), ^{t}BuOH:H_2O (1:1 v:v), rt, 90% yield, 2:1 d.r. 3. CF_3CO_2H (1.0 equiv), CH_2Cl_2, 23 °C, 93% yield. 4. $Fe(acac)_3$ (1.2 equiv), $PhSiH_3$ (3.0 equiv), EtOH, 65% yield. 5. LiHMDS (1.2 equiv), THF, 0 °C, then **12** (1.0 equiv); then KOH (10.0 equiv), EtOH, 80 °C, 68% yield and 11% recovered **11**. 6. MeLi (4.0 equiv), THF, 0 °C, 92% yield, >20:1 d.r. 7. $AlCl_3$ (1.1 equiv), CH_2Cl_2, –15 °C, 44% yield, **15:16** = 3:1. 8. $TiCl_4$ (1.3 equiv), Et_3SiH (2.0 equiv), CH_2Cl_2, 0 °C, 88% yield. 8'. $AlMe_3$ (5.0 equiv), diisopropylamine (6.0 equiv), 1,2-dichloroethane, 80 °C, 92% yield. 9'. *m*-chloroperbenzoic acid (1.2 equiv), THF, 0 °C; then Dess-Martin periodinane (4.0 equiv), THF, 23 °C, 56% yield. 10'. HBr (0.5 equiv), THF, 23 °C to 40 °C, 69% yield. 11'. $NaBH_4$ (4.0 equiv), THF:MeOH (1:1 v:v), 0 °C, 96% yield, >20:1 d.r. 12'. $TiCl_4$ (3.0 equiv), Et_3SiH (3.0 equiv), CH_2Cl_2, 0 °C, 63% yield. 13'. $Cu(OAc)_2$ (2.0 equiv), AgTFA (2.0 equiv), $Pd(OAc)_2$ (30 mol%), 2-methyl-2-butene (30.0 equiv), MeCN, 32% yield.

We began our synthesis with the thermodynamic alkylation of known Wieland-Miescher-ketone derivative **7**, which was prepared in one step from commercial materials by Robinson annulation[24] with homoprenyl iodide **8**, resulting in formation of diketone **9** with a diastereomeric ratio (d.r.) of 3:1 (Fig. 3). Dihydroxylation of the more exposed alkene gave a separable mixture of diastereomers (2:1 d.r.), the major of which could be ketalized with TFA to provide **10**. Thermodynamic reduction using hydrogen atom transfer conditions with $Fe(acac)_3$ and $PhSiH_3$ afforded tricycle **11** in 65% yield[25,26] as a single detectable diastereomer as confirmed by X-ray crystallography. The indole moiety was introduced through alkylation of **11** with tosyl-protected indole iodide **12**, which was deprotected *in situ* by the addition of ethanolic KOH. The key precursor to the biomimetic transformation was obtained through MeLi addition to **13**, yielding the axial tertiary alcohol **14**, as confirmed by X-ray crystallography. The C-ring cyclization to form the cyclopentane ring was found to proceed under a limited set of reaction conditions with the optimal result obtained upon treatment with $AlCl_3$ in CH_2Cl_2. A mixture of the congeners **15** and **16** was isolated in 44% yield in a ratio of 3:1, favoring the methyl migration product **15**. It could be speculated that this methyl migration occurs via a concerted pathway, because the computational results suggested cyclization would be favored. The concerted migration may be favorable due to the proximal orientation of the migrating methyl group, as well as the release of *syn*-pentane strain that could occur as a result. Additionally, related 1,2-methyl migrations have been suggested to proceed via concerted pathways[20]. In order to modulate the product distribution, synthesis of the equatorial alcohol and derivatives with other leaving groups was attempted, but these intermediates could not be accessed due to inherent steric biases and functional group compatibility issues.

As a control experiment to validate the computational analysis, we tested the optimal cyclization conditions on the axial alcohol precursor to **4B**, which formed the methyl shift product *iso*-**20** in low yield while paspaline A could not be detected. This result is consistent with the prediction based on $\Delta\Delta G^{\ddagger}$ that cyclization would be more efficient with **4C** rather than **4B** or **4A**.

To complete the synthesis of paspaline A, cyclization product **16** was subjected to ionic reduction conditions with $TiCl_4$ and Et_3SiH to give paspaline A (**1**) in 9 steps from commercially available starting materials, which is approximately one third the number of steps as either the landmark 25-step Smith[7] or 27-step Johnson[9] syntheses.

In the formation of compound **15**, the elimination following the methyl migration proceeded directly to the undesired tetrasubstituted alkene. In order to transpose the olefin to the desired position, enol ether **17** was accessed in 92% yield by treatment of **15** with (*i*-Pr_2N)$AlMe_2$[27]. A Rubottom oxidation of the enol ether, followed by treatment with Dess-Martin periodinane, provided a β,γ-unsaturated ketone. Isomerization of the alkene was then elicited through treatment with HBr to give **18** in 69% yield. Removal of the carbonyl functionality could be effected through reduction of the enone to the allylic alcohol **19**, followed by simultaneous ionic reduction of both the allylic alcohol and ketal to give **20**. Subsequent *N*-reverse prenylation through a one-step protocol described by Baran and co-workers provided emindole PB (**2**) in greater yield (32%) than the conventional two-step propargylation and reduction approach (11%)[28]. The material obtained by this route produced spectroscopic data which agreed with that of Kawai and co-workers, and the structure was confirmed through comparative DFT NMR calculations and X-ray crystallography of the penultimate intermediate **20**.

DFT augmented retrosynthesis can assist in selecting substrates to be prioritized for experimentation and significantly streamline synthetic workflows. Here, we have demonstrated its utility in developing the synthesis of paspaline A in approximately one third the steps of previous routes[7,9], as well as the first synthesis of emindole PB. The limitations and capabilities of this computational approach are still unexplored, but further studies in predictive computational evaluation and experimentation will define the boundaries of this framework.

References

---

1. E. J. Corey, X.-M. Cheng, *The Logic of Chemical Synthesis* (Wiley, New York, 1995).
2. E. J. Corey, W. T. Wipke, *Science,* **166,** 178–192 (1969).
3. M. Kowalik, C. M. Gothard, A. M. Drews, N. A. Gothard, A. Weckiewicz, P. E. Fuller, B. A. Grzybowski, and K. J. M. Bishop, *Angew. Chem. Int. Ed.* **51**, 7928–7932 (2012).
4. C. J. Marth, G. M. Gallego, T. P. Lebold, S. Kulyk, K. G. M. Kou, J. Qin, R. Lilien, R. Sarpong, *Nature* **528**, 493–498 (2015).
5. M. Elkin, T. R. Newhouse, *Chem. Soc. Rev.* **47**, 7830–7844 (2018).

6 D. J. Tantillo, *Chem. Soc. Rev.* **47**, 7854–7850 (2018).

7. A. B. Smith, R. Mewshaw, *J. Am. Chem. Soc.* **107**, 1769–1771 (1985).

8. A. B. Smith, T. L. Leenay, *J. Am. Chem. Soc.* **111**, 5761–5768 (1989).
9. R. J. Sharpe, J. S. Johnson, *J. Am. Chem. Soc.* **137**, 4968–4971 (2015).
10. D. T. George, E. J. Kuenstner, S. V. Pronin, *J. Am. Chem. Soc.* **137**, 15410–15413 (2015).
11. M. Enomoto, A. Morita, S. Kuwahara, *Angew. Chem. Int. Ed.* **51**, 12833–12836 (2012).
12. J. P. Springer, J. Clardy, *Tetrahedron Lett.* **21**, 231–234 (1980).
13. K. Kawai, K. Nozawa, S. Nakajima, *J. Chem. Soc. Perkin Trans. 1* **0**, 1673–1674 (1994).
14. T. Hosoe, T. Itabashi, N. Kobayashi, S. Udagawa, K. Kawai, *Chem. Pharm. Bull.* **54**, 185–187 (2006).
15. A. A. Sallam, N. M. Ayoub, A. I. Foudah, C. R., Gissendanner, S. A. Meyer, K. A. El Sayed, *Eur. J. Med. Chem.* **70**, 594–606 (2013).
16. R. Nusse, H. Clevers, *Cell* **169**, 985–999 (2017).
17. K. Tagami, C. Liu, A. Minami, M. Noike, T. Isaka, S. Fueki, Y. Shichijo, H. Toshima, K. Gomi, T. Dairi, H. Oikawa, *J. Am. Chem. Soc.* **135**, 1260–1263 (2013).
18. M. C. Tang, H. C. Lin, D. Li, Y. Zou, J. Li, W. Xu, R. A. Cacho, M. E. Hillenmeyer, N. K. Garg, Y. Tang, *J. Am. Chem. Soc.* **137**, 13724–13727 (2015).
19. K. Nakanishi, M. Doi, Y. Usami, T. Amagata, K. Minoura, R. Tanaka, A. Numata, T. Yamada, *Tetrahedron* **69**, 4617–4623 (2013).
20. H. W. Whitlock, P. B. Reichardt, F. M. Silver, *J. Am. Chem. Soc.* **93**, 485–491 (1971).
21. S. P. T. Matsuda, W. K. Wilson, Q. Xiong, *Org. Biomol. Chem.* **4**, 530–543 (2006).
22. E. V. Anslyn, D. A. Dougherty, *Modern Physical Organic Chemistry* (University Science Books, 2006).
23. X. Xiong, D. Zhang, J. Li, Y. Sun, S. Zhou, M. Yang, H. Shao, A. Li, *Chem. Asian J.* **10**, 869–872 (2015).
24. K. K. Wan, K. Iwasaki, J. C. Umotoy, D. W. Wolan, R. A. Shenvi, *Angew. Chem. Int. Ed.* **54**, 2410–2415 (2015).
25. K. Iwasaki, K. K. Wan, A. Oppedisano, S. W. M. Crossley, R. A. Shenvi, *J. Am. Chem. Soc.* **136**, 1300–1303 (2014).
26. S. M. King, X. Ma, S. B. Herzon, *J. Am. Chem. Soc.* **136**, 6884–6887 (2014).
27. K. C. Nicolaou, S. Kubota, W. S. Li, *J. Chem. Soc. Chem. Commun.* **0**, 512–514 (1989).
28. M. R. Luzung, C. A. Lewis, P. S. Baran, *Angew. Chem. Int. Ed.* **48**, 7025–7029 (2009).

Acknowledgments: We gratefully acknowledge Dr. Brandon Mercado for X-ray crystallography of compounds **11**, **14**, and **20** Yingchuan Zhu for technical assistance.

Funding: Financial support for this work was provided by Yale University, the Sloan Foundation, the National Science Foundation (GRFP to J.E.Z.). We gratefully acknowledge the National Science Foundation for financial support in the establishment of the Yale University High Performance Computing (HPC) Center (CNS 08-21132).