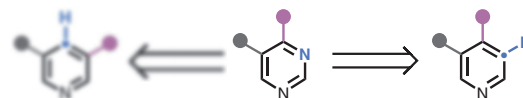


Retrosynthetic Simplicity

Mark D. Levin* 

Department of Chemistry, University of Chicago, Chicago, IL
60637, USA
marklevin@uchicago.edu



*some disconnections
are easier to see*

Received: 08.10.2023

Accepted after revision: 06.11.2023

Published online: 09.11.2023 (Version of Record)

DOI: 10.1055/s-0043-1763623; Art ID: ST-2023-10-0456-SP

Abstract Retrosynthetic simplicity is introduced as a metric by which methods can be evaluated. An argument in favor of reactions which are retrosynthetically simple is put forward, and recent examples in the context of skeletal editing from my own laboratory as well as contributions from others are analyzed critically through this lens.

- 1 Introduction
- 2 Example 1: Quinoline to Quinazoline (One Product, Two Starting Materials)
- 3 Example 2: Benzene to Pyridine (Two Products, One Starting Material)
- 4 Caveats, Counterarguments, and Conclusions

Key words skeletal editing, retrosynthesis, simplicity, heterocycles, methodology



Mark Levin was born and raised in Cleveland, OH in a Russian-Jewish immigrant family. In 2012, he earned a BSc in chemistry from the University of Rochester conducting undergraduate research with Professor Alison J. Frontier as well as with Henning Hopf at TU Braunschweig. He completed his PhD with Professor F. Dean Toste at the University of California, Berkeley in 2017 studying the redox chemistry of gold and conducted postdoctoral research in enantioselective catalysis with Professor Eric N. Jacobsen at Harvard University. In 2019, Mark joined the University of Chicago as an assistant professor of chemistry and was promoted to associate professor with tenure in 2022. His research interests include the development of synthetic methodology and study of their mechanisms, with a specific focus on single-atom skeletal editing.

1 Introduction

When planning a synthesis, why is it that our minds reach for some reactions and not others? This is a central question for those of us interested in developing new reactions, because it influences whether your work will get used. There are countless reasons to develop a chemical transformation that have nothing to do with an end user, but I suspect many methodologists share my hope that the reagents, catalysts, and ideas we put out into the world will end up in someone else's flask.

There are some obvious metrics that authors (and reviewers) spend their time considering – How easy is the reaction to run? Can I buy all the reagents or make them easily? How broad is the scope? (i.e., will it work on *my* molecule?) Is it selective? Is it safe? These are clearly important components for adoption, but I would contend they are not a sufficient set of criteria. By inspection, one can ascertain that there are plenty of good reactions by this rubric that do not enjoy a broad userbase, as most papers do not generate

a widely adopted method. Why? Visibility almost certainly has something to do with this – many good reactions are published in more specialized settings that don't attract as large of a readership, at least initially (though the evidence on the effect of visibility is mixed^{1,2}). The structure of our introductory courses likewise makes some impact, but again inspection reveals that we don't hew exclusively to pedagogically privileged reactions when conducting syntheses.

So, what else might be missing? I propose an underappreciated criterion, the surprise of which is spoiled by this article's title: retrosynthetic simplicity. While we like to think of ourselves as artistic and sophisticated planners of syntheses, and while we as a community rightfully celebrate insightful and surprising disconnections, when planning a routine synthesis (e.g., preparation of a substrate), this self-image does not persist. Rather, we avoid expending

unnecessary mental energy taking our molecules apart *despite* our appreciation for elegant flourishes, gravitating instead towards disconnections that are ‘easy to see’. Compare for instance the ease with which a Suzuki coupling opportunity can be identified compared to a [3,3]-sigmatropic rearrangement (Figure 1).^{3,4}

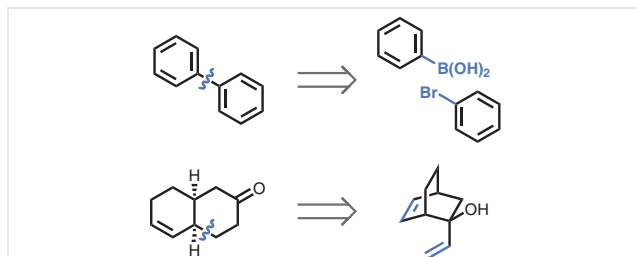


Figure 1 Which of these retrosynthetic disconnections is easier to identify? Which one are you more likely to come up with? Does that impact how often each is used?

The proliferation of cross-coupling as a workhorse synthetic technology^{5–8} can be seen as a consequence of its incorporation of all the above desiderata – generality, modularity, accessibility, ease of operation – while *also* exemplifying simplicity in its associated retrosynthesis. Indeed, for better or for worse, we have adopted the nomenclature of cross-coupling so widely that a large percentage of new methods (whether cross-coupling or not) are now named by the kind of bond they form; this ‘cross-coupling-ization’ of how reactions are categorized reveals how deeply rooted cross-coupling’s retrosynthetic logic is in our collective consciousness.

Simplicity and complexity in chemistry are very difficult to define unambiguously,^{9–12} so rather than wade into that particular morass I will pose a somewhat impertinent definition of retrosynthetic simplicity: retrosynthetically simple reactions are those where you do not have to be a genius, nor indeed work very hard, to identify where they can be applied. Reducing to absurdity, cross-coupling is popular, in part, because even someone who doesn’t understand bond-line notation can be quickly and easily taught to identify when one might be appropriate.

These thoughts were a motivating force behind the choice for my laboratory to pursue ‘single-atom’ level modifications rather than more dramatic multiatom reactions when we first set out to make our mark in skeletal editing.^{13–15} Indeed, though the broader ‘molecular editing’ buzzword is (taken literally) a synonym for ‘chemical reaction’, – and while the phrase itself has become unmoored from its original conception in natural-product-based medicinal chemistry^{16–18} – the idea that editing should be reserved to describe small, simple changes is so intuitive that it rarely needs explicit demarcation. (If I rewrite a document beyond recognition, I have not edited it.) However, even among single-atom-editing reactions and sequences, there remains some variance in the retrosynthetic simplici-

ty that one is confronted with, depending on the details. For two illustrative examples, I focus below on recently reported examples of C-to-N replacement reactions from my own laboratory, contrasted with alternative approaches that achieve similar end results.

2 Example 1: Quinoline to Quinazoline (One Product, Two Starting Materials)

Recently, my laboratory reported a transformation that enables the conversion of quinoline *N*-oxides into quinazolines by replacement of the C3 carbon with an ammonia nucleophile, eliminating a carboxylate leaving group.¹⁹ However, we had previously also reported that acidolysis of the intermediate benzoxazepine involved in this transformation leads instead to the formation of the C2 carbon deletion product – an indole.²⁰ One could imagine combining this carbon deletion with the Morandi laboratory’s nitrogen insertion into indoles, also affording a quinazoline.²¹

Though it is tempting to treat these as equivalent, rival skeletal-editing approaches to quinazolines (perhaps differentiated by their step counts), they differ critically in their retrosynthetic logic (Figure 2). Whereas the direct C3 replacement can be envisioned starting from a quinoline wherein your mind’s eye replaces the nitrogen of the quinazoline with a methine, the latter route, by virtue of the mismatched insertion and deletion, requires one to begin with the 2-substituent of the quinazoline in the 3-position of the quinoline. All else equal (e.g., for a given set of quinolines the two isomers may not be equally accessible), when planning this synthesis, it is more natural to reach for the starting material whose substitution pattern most closely matches the target.

3 Example 2: Benzene to Pyridine (Two Products, One Starting Material)

Another transformation recently developed by my laboratory involves the conversion of aryl azides into pyridines through a two-step protocol involving initial azepine formation via an aryl nitrene, followed by oxidative spirocyclization and elimination of a carbene.²² Because the carbon deletion and the nitrogen insertion are matched in their regioselectivity, the resulting pyridine forms with nitrogen replacement of the former *ipso* carbon of the aryl azide. This is true regardless of the selectivity in the initial azepine formation, with both potential isomers converging to the same product and thus avoiding any skeletal rotation. An alternative skeletal-editing protocol involving azepine formation was reported by the Burns laboratory, in which singlet oxygen instead promotes deletion of the former *meta* carbon, leading to an aminopyridine.²³

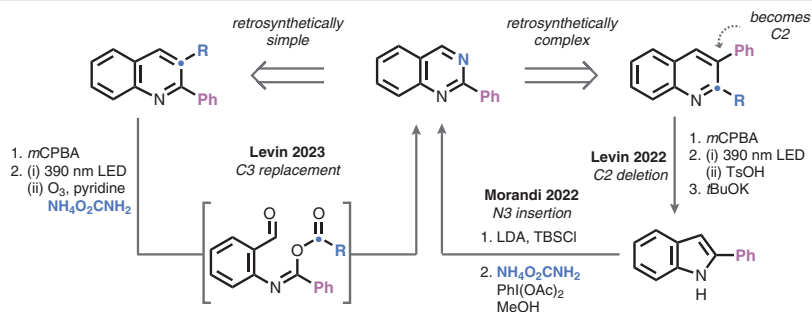


Figure 2 Two skeletal-editing approaches to the same quinazoline from different quinoline starting materials

While here it is tempting to view these merely as regiochemically divergent treatments of the same starting material, the retrosynthetic logic is again substantially different. As shown in Figure 3 for an example substrate, thinking backwards from the aminopyridine requires a rotation of the skeleton, whereas no such rotation is necessary to consider in the *ipso* deletion pathway. To be clear, there may be instances where the *meta* deletion approach is preferable for any number of reasons (wanting to make an aminopyridine, inability to access the isomeric azide, etc.), but it should also be clear from this analysis that the more natural retrosynthesis is the one that avoids perturbation of the remaining skeleton.

4 Caveats, Counterarguments, and Conclusions

The goal of this analysis is not to provide a measuring stick by which all reactions or syntheses should be evaluated, as there are important and useful venues for transformations that defy simplistic retrosynthesis, and these should continue to be developed and celebrated. Likewise, it can be the case that a simpler retrosynthesis leads to a less

accessible starting material, such that it will sometimes be preferable to take the less obvious route for practical reasons. It should also be noted that reactions with wildly complex mechanisms can still correspond to reactions whose retrosynthesis is easy to see, as the details of the mechanism do not *need* to be considered in synthetic planning. (Though they can of course help determine whether your target is in the scope, one can still make successful predictions even with incorrect mechanisms.) Finally, the notion of retrosynthetic simplicity should be understood as a continuum rather than as a dichotomy.

An anonymous reviewer of the initial version of this manuscript posed two additional challenges that merit some discussion. The first of these involves the relation of the present discussion to Corey's concept of 'structurally simplifying' transforms.²⁴ Despite the commonality of simplicity as a metric, the two ideas are orthogonal because retrosynthetic simplicity is typically maintained when the reaction is reversed while simplifying transforms are not. For example, if you argue that the disconnection of a quinazoline to a quinoline decreases functional group content, then the reverse transform is still retrosynthetically simple but not structurally simplifying. As such, a reaction can be retrosynthetically simple without offering a struc-

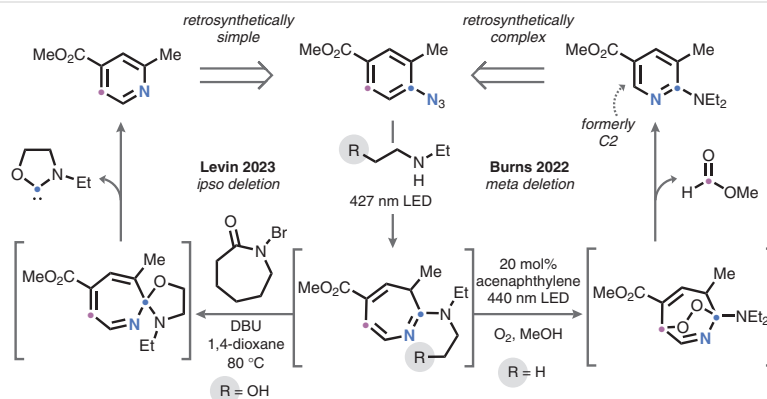


Figure 3 Two skeletal editing approaches to different pyridines from the same aryl azide starting material

turally simplifying transform. The converse is also true – as mentioned above, examples in natural product synthesis abound.

On that note, the second challenge concerns truly weighty synthetic plans (i.e., natural product syntheses), wherein one rarely employs the first route that comes to mind – to quote the reviewer, ‘why would one be so lazy?’. I don’t disagree with this point, but I think context matters. There is a large volume of ‘routine’ synthesis that does not rise to the level of planning scrutiny that total synthesis campaigns entail. Indeed, many chemists employ organic synthesis without thinking of themselves (primarily) as organic chemists, and they too are a potential audience for new methods. A related criticism is the recognition of the role of software (e.g., SciFinder or more recently automated retrosynthesis) in route planning – ‘it doesn’t matter if [a human] can see a transform or not’.^{25,26} However, the influence of software is not in conflict with the ideas presented here. Unrestrained searches typically yield far more results than can be reasonably parsed, requiring some user involvement in prioritization. This injection of the planner’s psychology brings retrosynthetic simplicity into play, for better or for worse. Automated retrosynthesis (artificially intelligent or otherwise) will likewise reflect its training dataset to some extent, which is going to encode the biases of the chemists who run some reactions more often than others, regardless of the reasons for that asymmetry.²⁷ If you want these algorithms to recommend your new method, you’ll need to convince (some) chemists to use it first.

With these caveats and arguments in mind, it is my hope that in the design of new reactions, in addition to the laundry list of criteria that are already used to assess their worth,^{28–31} that the ease with which a potential end-user can identify their application is considered as well – not in place of the other criteria, but when appropriate, in addition. As the analyses presented here hopefully demonstrate, retrosynthetic simplicity goes beyond aesthetic appeal and offers a guide for developing reactions that others can naturally conceive of applying in their own syntheses. Put another way, the beauty inherent in simplicity is often appreciated,³² but there is a hidden utility in simplicity, too.

Conflict of Interest

The authors declare no conflict of interest.

Funding Information

This manuscript highlights work, published elsewhere, that was supported by the National Institute of General Medical Sciences (R35 GM142768).

Acknowledgment

The ideas presented here were developed through many conversations with members of my research laboratory over the past four

years, and I am indebted to all of them for their contributions to this work. I also wish to thank Andy McNally (Colorado State), Jake Essman (Harvard), Zach Wickens (Wisconsin), and Alison Wendlandt (MIT), and one of the anonymous reviewers for insightful comments during the preparation of this manuscript.

References

- (1) Chan, H. F.; Önder, A. S.; Schweitzer, S.; Torgler, B. *Econ. Lett.* **2023**, *231*, 111270.
- (2) Branch, T. A.; Côté, I. M.; David, S. R.; Drew, J. A.; LaRue, M.; Márquez, M. C.; Parsons, E. C. M.; Rabaiotti, D.; Shiffman, D.; Steen, D. A.; Wild, A. L. *bioRxiv* **2023** DOI: preprint; 10.1101/2023.09.17.558161
- (3) Miyaura, N.; Suzuki, A. *Chem. Rev.* **1995**, *95*, 2457.
- (4) Evans, D. A.; Golob, A. M. *J. Am. Chem. Soc.* **1975**, *97*, 4765.
- (5) Ghosez, L. *Isr. J. Chem.* **2018**, *58*, 127.
- (6) Brown, D. G.; Boström, J. *J. Med. Chem.* **2016**, *59*, 4443.
- (7) Roughley, S. D.; Jordan, A. M. *J. Med. Chem.* **2011**, *54*, 3451.
- (8) Wang, Y.; Haight, I.; Gupta, R.; Vasudevan, A. *J. Med. Chem.* **2021**, *64*, 17115.
- (9) Bertz, S. H. *J. Am. Chem. Soc.* **1982**, *104*, 5801.
- (10) Whitesides, G. M.; Ismagilov, R. F. *Science* **1999**, *284*, 89.
- (11) Fuchs, P. L. *Tetrahedron* **2001**, *57*, 6855.
- (12) Woo, S.; Shenvi, R. A. *Acc. Chem. Res.* **2021**, *54*, 1157.
- (13) Jurczyk, J.; Woo, J.; Kim, S. F.; Dherange, B. D.; Sarpong, R.; Levin, M. D. *Nat. Synth.* **2022**, *1*, 352.
- (14) Boger, D. L. *J. Org. Chem.* **2017**, *82*, 11961.
- (15) Kennedy, S. H.; Dherange, B. D.; Berger, K. J.; Levin, M. D. *Nature* **2021**, *593*, 223.
- (16) Wilson, R. M.; Danishefsky, S. J. *J. Org. Chem.* **2006**, *71*, 8329.
- (17) Szpilman, A. M.; Carreira, E. M. *Angew. Chem. Int. Ed.* **2010**, *49*, 9592.
- (18) Fan, Z.; Chen, X.; Tanaka, K.; Park, H. S.; Lam, N. Y. S.; Wong, J. J.; Houk, K. N.; Yu, J.-Q. *Nature* **2022**, *610*, 87.
- (19) Woo, J.; Stein, C.; Christian, A. H.; Levin, M. D. *Nature* **2023**, *623*, 77.
- (20) Woo, J.; Christian, A. H.; Burgess, S. A.; Yuan, J.; Mansoor, U. F.; Levin, M. D. *Science* **2022**, *376*, 527.
- (21) Reisenbauer, J. C.; Green, O.; Franchino, A.; Finkelstein, P.; Morandi, B. *Science* **2022**, *377*, 1104.
- (22) Pearson, T. J.; Shimazumi, R.; Driscoll, J. L.; Dherange, B. D.; Park, D.-I.; Levin, M. D. *Science* **2023**, *381*, 1474.
- (23) Patel, S. C.; Burns, N. Z. *J. Am. Chem. Soc.* **2022**, *144*, 17797.
- (24) Corey, E. J.; Cheng, X.-M. *The Logic of Chemical Synthesis*. Hoboken Wiley. **1995**.
- (25) Mikulak-Klucznik, B.; Gołębiowska, P.; Bayly, A. A.; Popik, O.; Klucznik, T.; Szymkuć, S.; Gajewska, E. P.; Dittwald, P.; Staszewska-Krajewska, O.; Beker, W.; Badowski, T.; Scheidt, K. A.; Molga, K.; Mlynarski, J.; Mrksich, M.; Grzybowski, B. A. *Nature* **2020**, *588*, 83.
- (26) Lin, Y.; Zhang, R.; Wang, D.; Cernak, T. *Science* **2023**, *379*, 453.
- (27) Beker, W.; Roszak, R.; Wołos, A.; Angello, N. H.; Rathore, V.; Burke, M. D.; Grzybowski, B. A. *J. Am. Chem. Soc.* **2022**, *144*, 4819.
- (28) Trost, B. M. *Science* **1983**, *219*, 245.
- (29) Trost, B. *Science* **1991**, *254*, 1471.
- (30) Gaich, T.; Baran, P. S. *J. Org. Chem.* **2010**, *75*, 4657.
- (31) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004.
- (32) von Goethe, J. W. *Goethe's Opinions on the World, Mankind, Literature, Science, and Art*. London J. W. Parker and Son. **1853**.