Chapter 8
Ranking of Synthesis Plans

Abstract Hendrickson's definition of the "ideal synthesis" serves as a benchmark to assess synthesis plans. Criteria such as convergency, increase in complexity, and robustness are presented to rank synthesis plans and to pinpoint weaknesses therein.

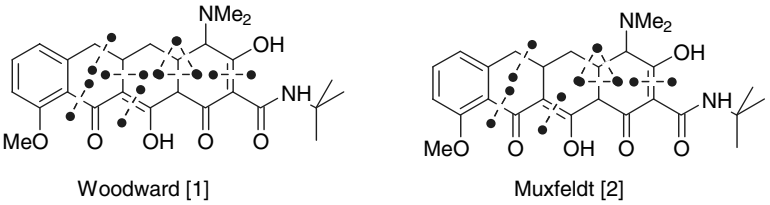
Any ranking of plans for the synthesis of a given target compound depends on benchmarks which must be defined. Possible criteria may be

- the shortest route (time involved),
- the cheapest route (cost of materials),
- the novelty of the route (patentability),
- the greenest route (avoidance of problematic waste),
- the healthiest route (avoidance of toxic intermediates and side products),
- the most reliable route (lowest risk approach).

Aside from these external criteria, ranking of synthesis proposals could also follow systematic criteria, e.g., the step count. A synthesis that reaches the target in fewer steps than another one is considered superior. Every synthesis consists of obligatory steps, i.e., those by which the skeleton is made. When focusing on this aspect, the bond-set would give a lower limit to the number of steps involved in a projected synthesis, because any refunctionalization steps and protecting group management steps count in addition to the skeleton forming steps. Because of this, the bond-set does not reveal too much about a step count and the quality of a projected synthesis. For instance, the differences between the bond-sets of Woodward's [1] and Muxfeldt's [2] tetracycline syntheses are minimal (Scheme 8.1). Comparison of the bond-sets does not reveal that in Muxfeldt's synthesis three bonds are formed in one operation rendering this synthesis significantly shorter—22 steps in Woodward's synthesis versus 17 in Muxfeldt's.

Azurity Ex. 1029
IPR Petition – USP 12,128,039

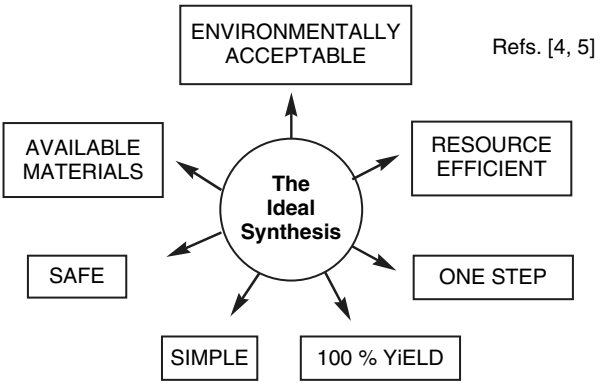
R.W. Hoffmann, *Elements of Synthesis Planning*,
DOI 10.1007/978-3-540-79220-8_8, © Springer-Verlag Berlin Heidelberg 2009



Scheme 8.1 Bond-sets of Woodward's and Muxfeld's syntheses of tetracycline

A comparison of the number of bonds in the bond-set with the total step count reveals that refunctionalization operations constitute the lion's share of the steps. When the step count is considered as a decisive criterion to judge the quality of a synthesis, then it becomes time to discuss some comments regarding the "ideal synthesis" (Scheme 8.2).

An "ideal synthesis" has been postulated by Turner [3] and by Wender [4, 5] as a one-pot reaction in which all starting materials are mixed, leading directly to the final product.



Scheme 8.2 Wender's comments regarding an ideal synthesis

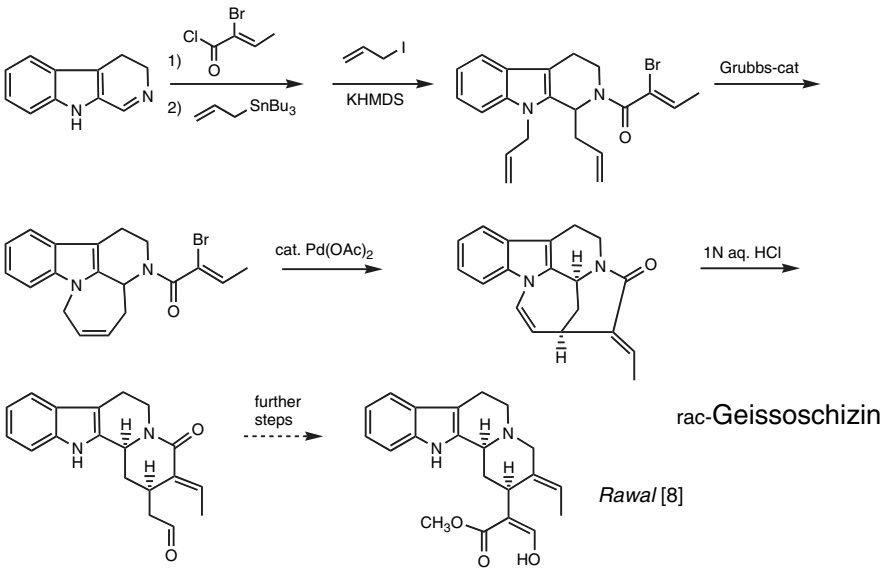
"An ideal synthesis is generally regarded as one in which the target molecule is prepared from readily available, inexpensive starting materials in one simple, safe, environmentally acceptable, and resource-efficient operation that proceeds quickly and in quantitative yield."

Given the present imperfection of one-pot syntheses [6], this definition of the "ideal synthesis" is utopian, at least for in vitro chemical synthesis. One has to admit, though, that the in vivo biosynthesis of natural products in a living cell by and large fulfills the criterion of a so-defined ideal synthesis.

I would prefer a more realistic definition of the ideal synthesis, one from which guidelines for the planning of syntheses can be derived. This is the definition by Hendrickson [7]:

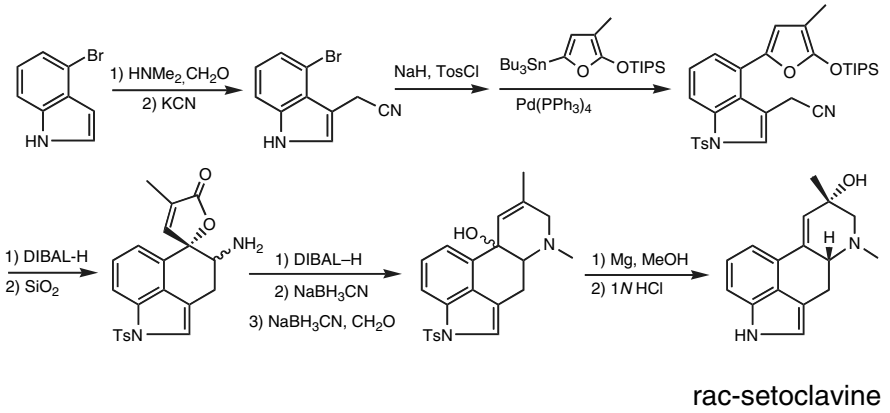
"Ideal Synthesis: The ideal synthesis creates a complex molecule … in a sequence of only construction reactions involving no intermediary refunctionalizations, leading directly to the target, not only its skeleton but also its correctly placed functionality."

The reasoning of Hendrickson [7] is that only the skeletal bond-forming steps are obligatory in a synthesis. The superfluous refunctionalization steps could be dispensed with, once one succeeds at generating in the skeletal bond-forming reactions exactly that functionality which is present in the target or which is required to carry out the next skeletal bond-forming reaction. According to this reasoning, a numerical benchmark for the quality of a synthesis could be the ratio of the skeletal bond-forming reaction steps to that of the refunctionalization steps. For the tetracycline syntheses shown in Scheme 8.1, the ratios come out to be $6/16 = 0.37$ (Woodward) and $5/12 = 0.41$ (Muxfeldt). There are indeed syntheses reaching much higher numerical values for this ratio, e.g., 2.5 for Rawal's synthesis of geissoschicine [8] (Scheme 8.3).



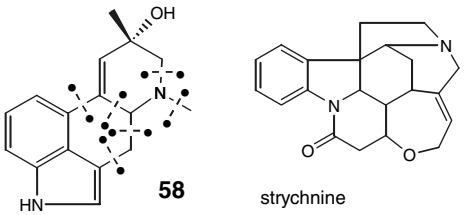
Scheme 8.3 Predominately skeleton bond-forming reactions in the geissoschicine synthesis by Rawal

A synthesis does not necessarily become good simply because it contains many skeletal bond-forming steps. This can be seen from an inspection of the setoclavine synthesis (ratio 1.67) [9] shown in Scheme 8.4.



Scheme 8.4 Predominately skeletal bond-forming reactions in setoclavine synthesis

The bond set of this setoclavine synthesis (**58,** Scheme 8.5) reveals that it is a piecemeal approach combining a multitude of small pieces (often C_1).

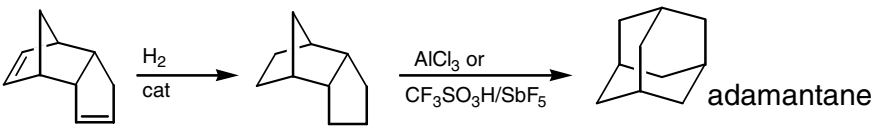


Scheme 8.5 Bond-set of the setoclavine synthesis in Scheme 8.4

At this point one could reach the nearly trivial conclusion that a synthesis is better, the shorter it is (the least number of bonds in the bond-set); the fewer the number of refunctionalization steps involved (oxidations, reductions); the fewer the number of protecting group steps; and the more it relies on reactions that form two or more skeletal bonds at a time. This accounts for the significant attention that is presently being given to tandem reactions [10] and reaction cascades.

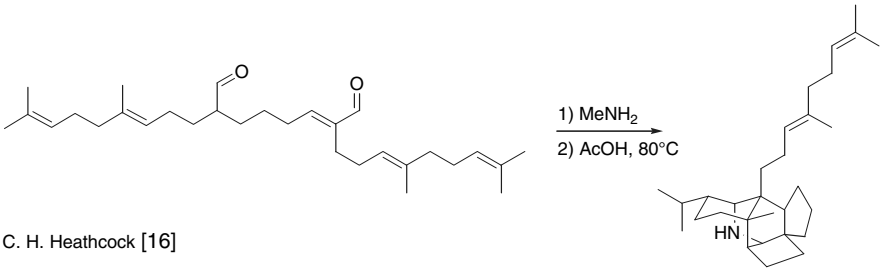
It is a cardinal principle of synthesis that it evolves from simple starting materials at the beginning to a complex structure in the end. During a synthesis sequence the complexity of the intermediates increases in a discontinuous manner. The *increase in complexity*—more precisely, how rapid

it is and at which point of the synthetic sequence it occurs—forms another important criterion to rank syntheses. The complexity of a target structure or of an intermediate is not, however, defined in a generally accepted manner. One could subscribe to the statement by Robinson [11] that strychnine (Scheme 8.5) is the most complex molecule known in relation to its size. When chemists speak of the complexity of a compound, they not only imply a complexity inherent in the structure, being topologically defined by the network of bonds. They also think about the difficulties of a synthesis, which may be a consequence of a lack of appropriate methods for synthesis [12]. Accordingly, adamantane (Scheme 8.6) was for a long time considered a rather complex target molecule. However, after a surprisingly simple synthesis of adamantane was realized by Schleyer [13], adamantane lost that status and is now a readily available commodity [14, 15]. Planning the synthesis of a compound like adamantane by a thermodynamics-driven equilibration process requires, however, completely different rules of synthesis planning than those presented in this text.



Scheme 8.6 Thermodynamics-driven formation of adamantane

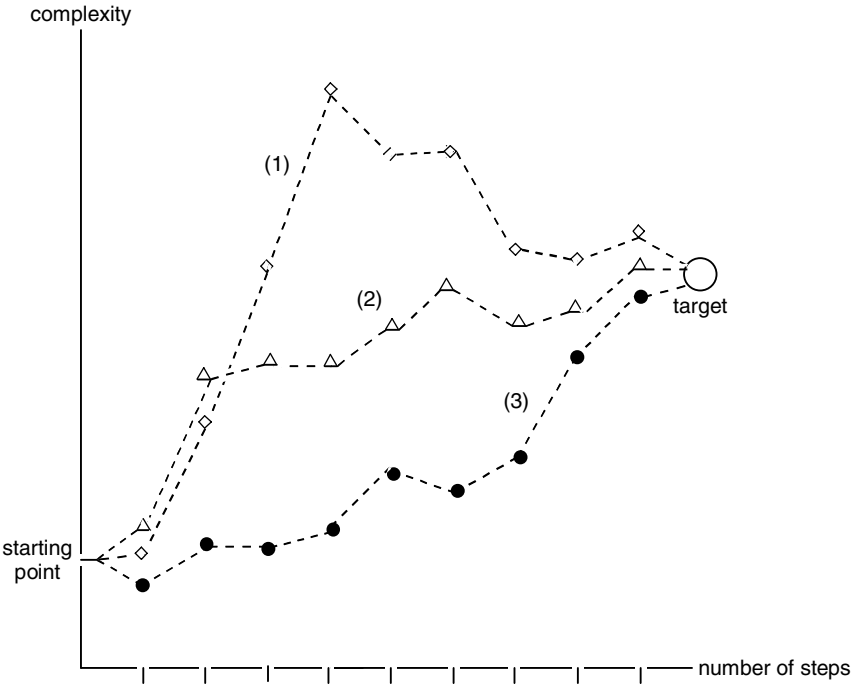
How is the complexity of a structure related to synthesis planning? One should start with the following truism: reactions with simple molecules proceed readily and in high yields. Complex molecules are frequently "touchier" and tend to form side products even in simple looking transformations. While this is certainly not a natural law, most chemists would attribute considerable truth to this statement. The consequence for planning a synthesis is this: keep the number of steps low, once your intermediates have become increasingly complex. A rapid increase in complexity early in a synthesis sequence frequently faces low yields in the (many) operations that follow. Sequences involving overbred molecular skeletons, i.e., those with a complexity higher than the following intermediates or the target structure, can be justified only when they allow a substantial reduction in the overall number of steps. Based on the above considerations, a late (exponential) increase of complexity in a synthesis sequence is most desirable. A most striking example is given in Scheme 8.7.



C. H. Heathcock [16]

Scheme 8.7 Rapid increase in complexity by a pentacyclization reaction

For analyses of synthesis sequences it will be of interest to quantify the complexity of both intermediates and final products. To this end, complexity indices have been developed which are based either on the topological complexity of the structures alone [12], or which, in addition, regard qualitatively the anticipated difficulties in a synthesis (intricacy indices) [17]. For a comparison of syntheses it is instructive to plot the "complexity" of the intermediates against the steps in the sequence. This results in plots such as the one shown in Scheme 8.8. A more detailed analysis may



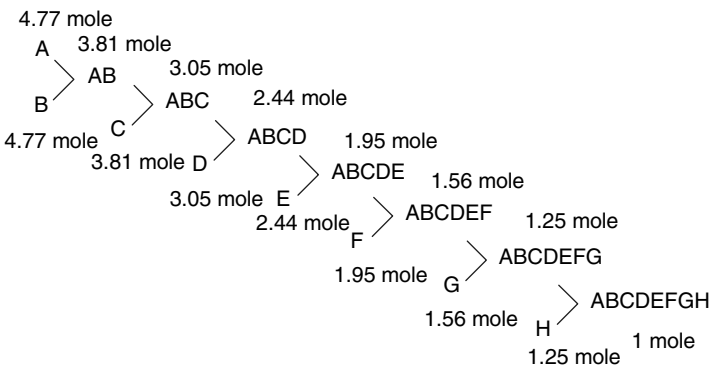
Scheme 8.8 Complexity trendlines of the intermediates in several (hypothetical) synthetic sequences for a given target

incorporate the structural similarity between the intermediates and the target [18] as well.

The charts arrived at in this manner have been analyzed by Bertz [19] in a paper strongly recommended for careful reading [cf. also 17]. Bertz adds the complexities of all intermediates along a synthesis to give the overall complexity. One is not surprised to find that the higher the overall complexity, the lower is the overall yield! This mirrors common experience. Again, the bottom-line is: avoid refunctionalization steps once the complexity of the intermediates has become high.

Of the synthetic sequences depicted as complexity trendlines in Scheme 8.8, sequence (3) appears to be the most attractive one, as it features a low overall complexity with a late-stage, rapid increase in complexity. In sequence (2), the complexity rises much too quickly—affording high overall complexity. Finally, sequence (1) features intermediates with complexity exceeding that of the target; hence, it should be given the lowest rank. Such an analysis allows one to pinpoint deficiencies in synthetic sequences, and to determine at which point in a synthetic sequence improvements would be of the greatest consequence [17].

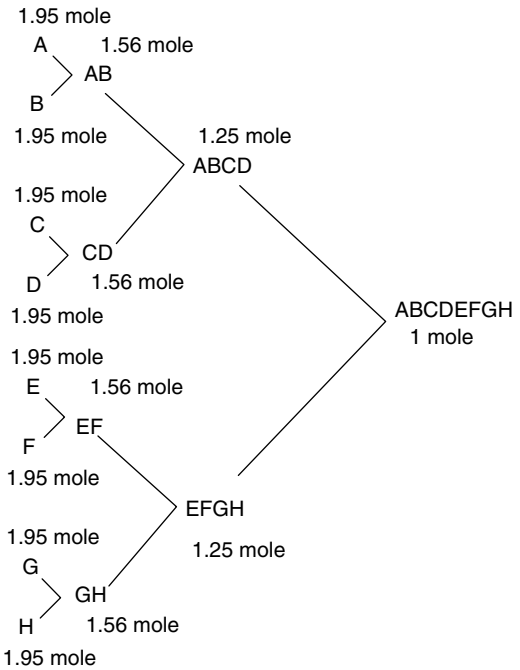
Such reasoning will favor syntheses that feature an exponential (= late) increase in complexity [20]. This can be influenced by the order in which the individual bonds in a bond-set are constructed. Here, the terms "linear" or "convergent" or "partially convergent" come into play. These terms were coined by Velluz as a consequence of his search for an optimal assembly of the steroid skeleton [21]. Convergent syntheses will afford a much better material balance than linear sequences, as soon as the yields in the individual steps fall below being quantitative. This point is illustrated in Schemes 8.9 and 8.10 for the coupling of eight building blocks (A through H) to yield a final target ABCDEFGH assuming an average yield of 80% for each coupling step.



Scheme 8.9 Input of material for a linear synthesis with 80% average yield per step

If each and every coupling reaction proceeded quantitatively, only one mole of each of the starting materials would be needed to generate one mole of the target compound. If the yields for the coupling steps fall to about 80%, one is forced into a battle of materials: a total of 24 moles (instead of eight!) of starting materials is required to generate just one mole of the target compound. This means that 16 out of 24 moles (= to 2/3) of the material is lost in the form of side products or other losses—that is, waste that must be handled. Likewise, use of auxiliary reagents, neutralizing agents, and so forth is equally uneconomical.

In a fully convergent synthesis the situation is not perfect either, yet much more favorable (Scheme 8.10).

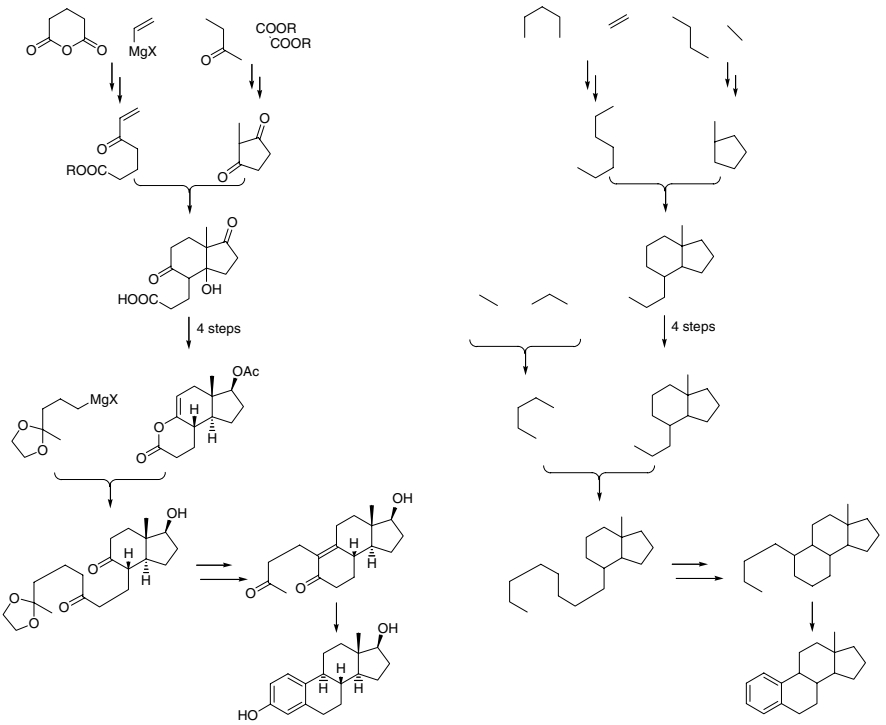


Scheme 8.10 Input of material for convergent synthesis with 80% average yield per step

In this case as well, one still needs a total of $(8 \times 1.95) = 15.6$ moles of starting materials to generate one mole of the target structure. The losses amount to $(15.6 - 8.0) = 7.6$ moles of building blocks, or almost half of the input materials. While remaining unsatisfactory, the ratio of output to input is much better in a convergent than in a linear synthesis. The more the average yields for the coupling steps approach the 100% limit, the smaller the difference gets between a linear and a convergent synthesis.

In this context the differences in "total yield" are frequently mentioned. "Total yield" refers to the yield of the final product ABCDEFGH calculated on the input of the initial first building block A. In a linear synthesis with average yields of 80% per step, the total yield for seven steps would amount to $0.8^7 = 21\%$. Since, in the hypothetical convergent synthesis, building block A would be carried through only three construction steps, the total yield would amount to $0.8^3 = 51\%$. Total yield is not a valid measure of the quality of a synthesis, because it measures the utilization of only the first building block, A. Utilization of a building block introduced later into the sequence will be higher (the yield based on that building block) as it is subjected to fewer overall yield-decreasing steps [22]. This is the basis for the advice to introduce "expensive" or otherwise "elaborate" building blocks as late as possible in a synthesis sequence!

Because of its historic importance for the development of the convergency principle, we summarize in Scheme 8.11 the estrone synthesis by Velluz [23].



Scheme 8.11 Overview of the estrone synthesis by Velluz (*left*, including functional groups; *right*, only molecular skeleton)

In the left half of Scheme 8.11 the full sequence of the estrone synthesis is given. In the right half merely the construction plan (i.e., the skeletal atoms and connectivities) [22, 24] is shown, highlighting the (partial) convergency [25].

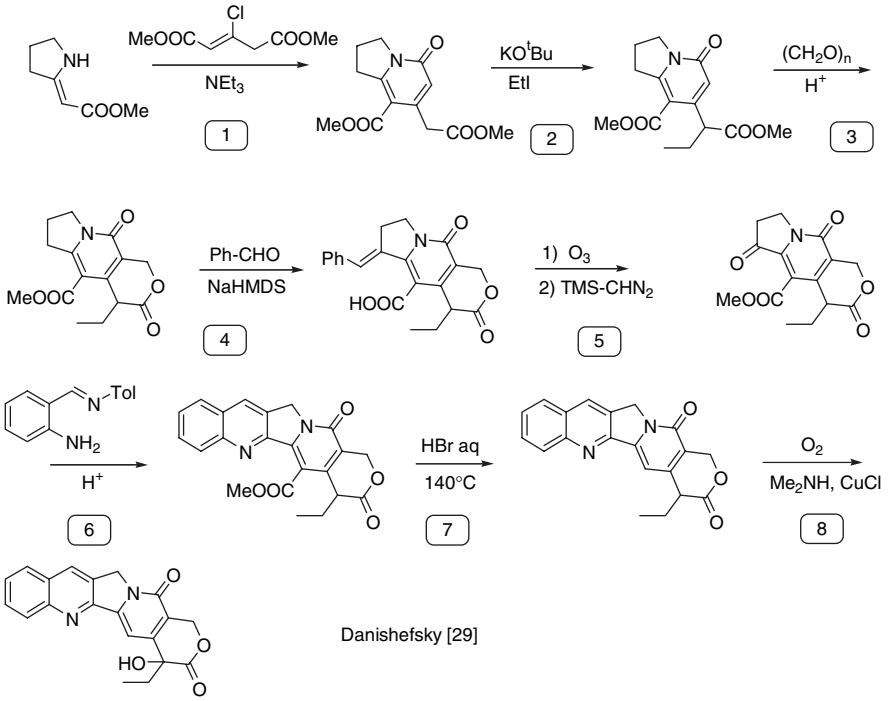
In summary, a good synthesis plan should be convergent, equivalent to a late-stage increase in complexity. The skeletal bond-forming steps should dominate in the step count. This means that refunctionalization steps and protecting group operations should be minimal, if they cannot be avoided entirely. In order to help meet these goals, consider the use of latent functionality.

Plans are subject to adverse events that may culminate in failure. This leads to the question: How *robust* is the plan to begin with? [26] One should check which and how many alternatives are available for each of the steps in a synthesis, in case the originally chosen variant cannot be employed. A robust synthesis plan will have at least one alternative for each step. Alternatives may include a change in the order in which bond-formations or refunctionalizations are to be carried out, or a longer sequence to reach a key intermediate. A plan that relies on a specific key step, which can be realized in only one conceivable manner, is a risky undertaking. One should evaluate how much of a synthesis tree falls apart when one specific step fails. What will be the consequence if, for instance, a six-membered ring cannot be accessed by a Diels-Alder cycloaddition as planned, and recourse has to be made to a Robinson-annulation? Which of the precursor molecules can still be used? Or does one have to begin again from a different set of starting materials? Which part of the protecting group scheme can be retained, once one is forced into an alternate route?

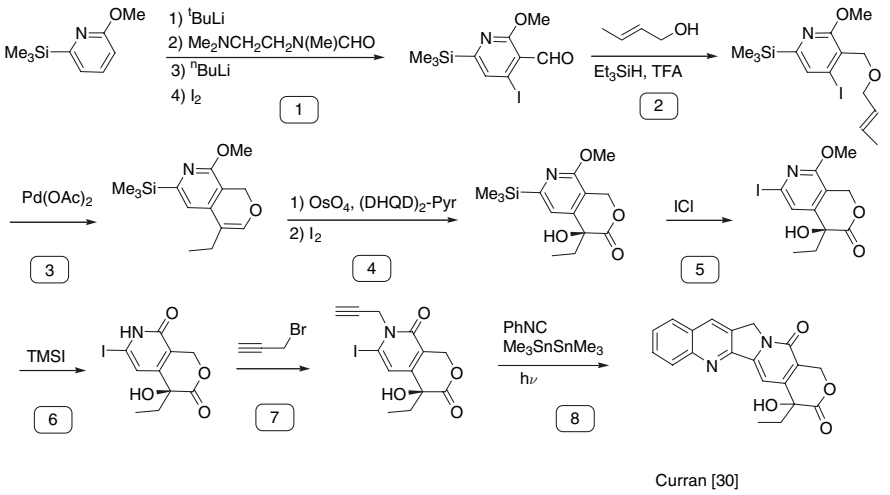
Few syntheses will reach the target in a manner that was originally planned (for example, see reference [27]). Modification of synthesis plans is called for every day. Hence, reasonable attention must be given to the robustness of the original plan. Stick to the principal advice [28]: get the most done in the fewest steps and with the highest yield.

Problems

8.1 Try to compare and evaluate the following two syntheses of camptothecin [29, 30]. Comment on the ratio of skeleton-building to refunctionalization steps, the increase in complexity, the degree of convergency, and the robustness of the synthesis plan.



Scheme 8.12 Danishefsky's camptothecin synthesis



Scheme 8.13 Curran's camptothecin synthesis

References

1. J. J. Korst, J. D. Johnston, K. Butler, E. J. Bianco, L. H. Conover, R. B. Woodward, *J. Am. Chem. Soc.* **1968**, *90*, 439–457.
2. H. Muxfeldt, W. Rogalski, *J. Am. Chem. Soc.* **1965**, *87*, 933–934.
3. S. Turner *The Design of Organic Synthesis*, Elsevier, Amsterdam, **1976**, p. 10.
4. P. A. Wender, *Chem. Rev.* **1996**, *96*, 1–2.
5. P. A. Wender, S. T. Handy, D. L. Wright, *Chem. Ind.* **1997**, 765.
6. S. J. Broadwater, S. L. Roth, K. E. Price, M. Kobaslija, D. T. McQuade, *Org. Biomol. Chem.* **2005**, *3*, 2899–2906.
7. J. B. Hendrickson, *J. Am. Chem. Soc.* **1975**, *97*, 5784–5800.
8. V. B. Birman, V. H. Rawal, *J. Org. Chem.* **1998**, *63*, 9146–9147.
9. S. Liras, C. L. Lynch, A. M. Fryer, B. T. Vu, S. F. Martin, *J. Am. Chem. Soc.* **2001**, *123*, 5918–5924.
10. L. F. Tietze, U. Beifuss, *Angew. Chem. Int. Ed. Engl.* **1993**. *32*, 131–163. (*Angew. Chem.* **1993**, *105*, 137–170).
11. R. Robinson, *Progr. Org. Chem.* **1952**, *1*, 1–21.
12. S. H. Bertz, *New J. Chem.* **2003**, *27*, 860–869.
13. R. C. Fort Jr., P. v. R. Schleyer, *Chem. Rev.* **1964**, *64*, 277–300.
14. M. A. McKervey, *Chem. Soc. Rev.* **1974**, *3*, 479–512.
15. G. A. Olah, D. Farooq, *J. Org. Chem.* **1986**, *51*, 5410–5413.
16. C. H. Heathcock, *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 665–681. (*Angew. Chem.* **1992**, *104*, 675–691).
17. P. L. Fuchs, *Tetrahedron* **2001**, *57*, 6855–6875.
18. M. Chanon, R. Barone, C. Baralotto, M. Julliard, J. B. Hendrickson, *Synthesis* **1998**, 1559–1583.
19. S. H. Bertz, *J. Am. Chem. Soc.* **1982**, *104*, 5801–5803.
20. S. H. Bertz, C. Rücker, G. Rücker, T. J. Sommer, *Eur. J. Org. Chem.* **2003**, 4737–4740.
21. L. Velluz, J. Valls, G. Nominé, *Angew. Chem., Int. Ed. Engl.* **1965**, *4*, 181–200. (*Angew. Chem.* **1965**, *77*, 185–205).
22. J. B. Hendrickson, *J. Am. Chem. Soc.* **1977**, *99*, 5439–5450.
23. L. Velluz, G. Nominé, G. Amiard, V. Torelli, J. Cérêde, *Compt. Rend. hebd. Acad. Sci.* **1963**, *257*, 3086–3088.
24. J. B. Hendrickson, E. Braun-Keller, G. A. Toczko, *Tetrahedron* **1981**, *37 Suppl.*, 359–370.
25. L. Velluz, J. Valls, J. Mathieu, *Angew. Chem., Int. Ed. Engl.* **1967**, *6*, 778–789. (*Angew. Chem.* **1967**, *79*, 774–785).
26. S. H. Bertz, *New J. Chem.* **2003**, *27*, 870–879.
27. R. Shen, C. T. Lin, E. J. Bowman, B. J. Bowman, J. A. Porco, *J. Am. Chem. Soc.* **2003**, *125*, 7889–7901.
28. R. E. Ireland *Organic Synthesis*, Prentice Hall, Englewood Cliffs, N.J., **1969**, p. 29.
29. W. Shen, C. A. Coburn, W. G. Bornmann, S. J. Danishefsky, *J. Org. Chem.* **1993**, *58*, 611–617.
30. D. P. Curran, S.-B. Ko, H. Josien, *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2683–2684. (*Angew. Chem.* **1995**, *107*, 2948–2950).