Perspectives on the application of green chemistry principles to total synthesis design

ABSTRACT
Green chemistry is reduced to a multi-variable optimization problem. Reaction mass efficiency (RME) is shown to be the key parameter that needs to be maximized in order for a valid claim that a reaction or synthesis is "green" can be made. The dependence of this parameter on various reaction parameters is illustrated for individual reactions. A brief summary is given of the key criteria and thresholds that link green chemistry principles to the analysis of individual reactions and to synthesis planning. A survey of literature plans for the total synthesis of the alkaloid (+)(S)-nicotine is used as an illustrative example. Key recommendations are made for the reporting of "green" syntheses.

INTRODUCTION

The new catch phrases "green chemistry" and "green engineering" are terms that refer to reaction optimization with respect to materials and energy usage, waste reduction from all sources, and overall cost minimization. Also included are minimization of toxicity and hazards, and maximization of safety practices. When these terms are applied to the performance of individual chemical reactions, chemical processes, and chemical synthesis plans they really deal with the same overriding issue: optimization of variables. Figure 1 summarizes the general "green chemistry" paradigm applied to chemical reactions so far implemented in the literature. Essentially all papers involve some combinatorial approach in which a benign medium is chosen along with one or more of the listed "green" strategies shown.

There have been a number of problems in getting green chemistry ideas accepted in chemical education and in the day-to-day practice of chemistry research. Among the most common are: (i) far too many coined names are given to obvious ideas, especially in the field of reaction metrics; (ii) few quantifiable connections are made between these ideas into a unified whole; (iii) inertia that green chemistry is perceived as "soft", not part of the rigorous physical sciences, precisely because of lack of mathematical rigor; (iv) confusing material and energy ideas when analyzing individual reactions or synthesis plans involving sets of chemical reactions; (v) diffuse meaning of sustainability with little mathematical rigor; (vi) prejudicing the training of synthetic organic chemists from the outset to avoid and regard any quantitative analysis and reasoning as not relevant to synthesis design; (vii) incorrect application of the traditional "overall yield" metric as a gauge of material efficiency of synthesis plans especially to convergent plans; and (viii) inability to connect and translate material and energy metrics to cost metrics, commonly referred to as the "bottom-line", that every professional in the chemical industry can understand, not just chemists and chemical engineers. In addition, there are disconnects and problems in current practices of carrying out total syntheses of target compounds, namely: (i) changing the reaction solvent to a benign one is considered sufficient to "green up" a chemical reaction, (ii) planning a synthesis with the constraint that a newly discovered reaction is used as the cornerstone reaction in a plan whether or not it is "green", or that use reactions that pay homage to past or present chemists who have made significant contributions to synthetic organic chemistry; (iii) not using multi-component reactions (MCR) (1 - 26) as part of the master plan in total synthesis design, but relying on the traditional linear approach; and (iv) the slow development of a library of stereoselective MCRs.

As noted in Anastas and Warner's classic text (27), "Green chemistry is the utilization of a set of principles that reduces or eliminates the use or generation of hazardous substances in the design, manufacture, and application of chemical products". The highest priority "green" principle is to carry out syntheses with the highest atom economy (28 – 30) while minimizing byproducts and the environmental impact factor (31 - 34). All of these ideas, though given new names, are in fact not new and are implemented routinely and intuitively by process chemists and chemical engineers in the chemical industry who operate under a code of good laboratory practice (GLP). *The bottom line concept for both green chemistry and green engineering is that ongoing efforts to improve reactions are directly connected to the optimization of key parameters that govern their performances.* It is obvious that some of these parameters will need to be maximized and others minimized. What is new is the recent introduction of a simple arithmetic description of optimization via the use of so called "green metrics" in a way that can be easily applied by any chemist to the widest range of chemical processes and synthesis plans regardless of complexity (35 - 39). The central metric by which material efficiencies of synthesis plans to a common target may be compared in a standard way is reaction mass efficiency (RME). For the purposes of satisfying the goals of green chemistry, this parameter needs to be maximized.

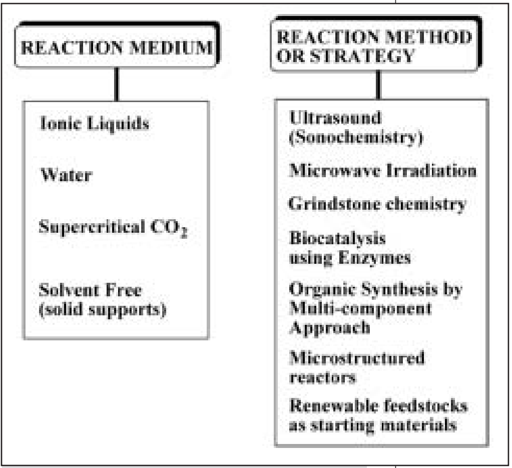


Figure 1. Summary of current green chemistry paradigm applied to organic reactions

This renaissance in thinking goes beyond the traditional "overall yield" parameter restricted to linear syntheses familiar to chemists. In this paper we highlight and illustrate the determination of RME for any individual reaction and compare various literature routes for the total synthesis of the alkaloid (+)(S)-nicotine.

REACTION MASS EFFICIENCY (RME)

The complete equation, derived elsewhere (35 – 39), for RME applicable to a single chemical reaction is given by

$$RME = (\varepsilon)(AE)\left(\frac{1}{SF}\right)(MRP) = (\varepsilon)(AE)\left(\frac{1}{SF}\right)\left(\frac{1}{1+\frac{\varepsilon(AE)[c+s+\omega]}{(SF)(m_P)}}\right) \quad (1)$$

where ε is the reaction yield (0 < ε < 1); AE is atom economy (0 < AE < 1); SF is the stoichiometric factor that takes into account the use of excess reagents (SF = 1 for stoichiometric reactions carried out with no excess reagents; SF >1, otherwise); MRP is the material recovery parameter that takes into account other materials used in the reaction and post-reaction phases (work-up and purification) such as solvents and washings for extractions ($0 < MRP < 1$); *c*, *s*, and ω are the masses of reaction catalyst, reaction solvent, and all other post-reaction materials respectively; and *mp* is the mass of the collected target product. Atom economy is given by the well known definition

$$AE = \frac{MW_{product}}{\sum MW_{reagents}} \qquad (2)$$

and the stoichiometric factor, by definition, is given by

$$SF = 1 + \frac{\sum mass_{excess\ reagents}}{\sum mass_{stoichiometric\ reagents}} = 1 + \frac{(AE)\sum mass_{excess\ reagents}}{theoretical\ mass_{product}}, \qquad (3)$$

The five parameters in eq 1 (RME, AE, ε, 1/SF, and MRP) may be displayed graphically in the form of a radial pentagon (see Figure 2) depicting a "materials usage footprint" so that chemists can recognize at once which of the four factors on the right hand side of eq 1 are contributing to an attenuation of RME. Each axis corresponding to one of the five parameters emanates from the centre and ranges in value between zero and one. The values of these parameters are depicted as dots and these are connected to form a pentagonal figure. The ideal "green" situation is depicted by a regular pentagon of unit radius where each parameter is equal to one. The less "green" a reaction is, the more the resultant pentagon is distorted toward the centre. These diagrams may therefore be used to compare the RME performances of different classes of reactions so that chemists can ascertain which reaction classes are inherently "green" and which are not by visual inspection. The degree of distortion of the radial pentagon from its regular ideal shape may be directly linked to which parameters are responsible for that distortion. These diagrams therefore inform chemists as to what to do if they wish to "green up" their experimental procedure, whether it is to reduce reaction solvent usage, cut down on unnecessary washes and extractions, avoid using excess reagents unless there is a chemical reason for doing so as in driving equilibria toward product, select lower mass reagents to effect improved atom economical performance, or optimize reaction yield by tweaking such parameters as reaction time, reaction temperature, reaction pressure, or the use of catalysts. A study on the use of these diagrams on undergraduate organic reactions has recently been described (39). The quantities AE and E_{mw} (Sheldon E-factor based on molecular weight), and RME and E_m (Sheldon E-factor based on mass) are related by simple expressions given by eq 4a and 4b which allow easy calculation of either parameter in each case once one of them is known. The derivations of eq 4a and 4b are linked to that of eq 1 and are based on the conservation of mass law connecting the molecular weights (or mass) of all input materials to the molecular weights (or mass) of all output materials.

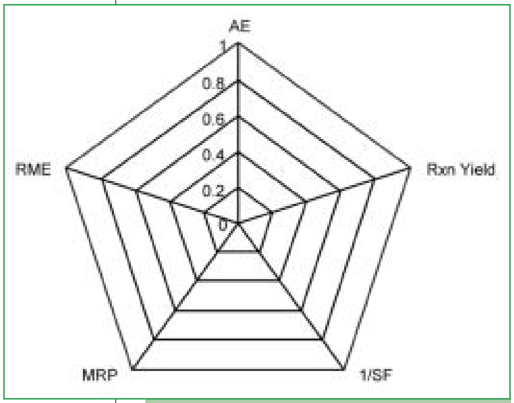


Figure 2. Radial pentagon representation of reaction mass efficiency (RME) showing dependence on four independent parameters given in equation (1)

$$RME = \frac{1}{1+E_m} \qquad (4a)$$

$$AE = \frac{1}{1+E_{mw}} \qquad (4b)$$

Equations (4a) and (4b) provide a means of assigning a set of criteria or mandatory requirements for claiming that a given reaction is "green"; namely, (i) AE > 61.8 percent so that AE > E_{mw}, (ii) RME > 61.8 percent so that RME > E_m, and (iii) reaction solvents and all post-reaction materials used in the work-up and purification stages must be reclaimed and/or eliminated (35). Figures 3 and 4 illustrate the number of named organic reactions in each class that meet this criterion and the progress of discovery of such reactions at the level of atom economy, respectively (37). Many reported claims of "green" reactions or syntheses in the literature fail these simple criteria.

APPLICATION OF RME TO SYNTHESIS PLANNING

When analyzing the material efficiencies of several synthesis plans to a given target it is sufficient to examine the kernel overall RME value based only on reaction yield and atom economy variables since these depend directly on the intrinsic chemical performances of the reactions in the plan. The calculation of this parameter for a given plan is greatly simplified if the plan is represented as a synthesis tree since it may be determined by inspection by a simple connect-the-dots approach. A recent paper introducing and describing the method showed that it can be applied to any synthesis plan of any degree of complexity from simple linear plans to highly convergent ones (38). This method may be used in conjunction with the standard retrosynthetic disconnect approach to weed out poor material

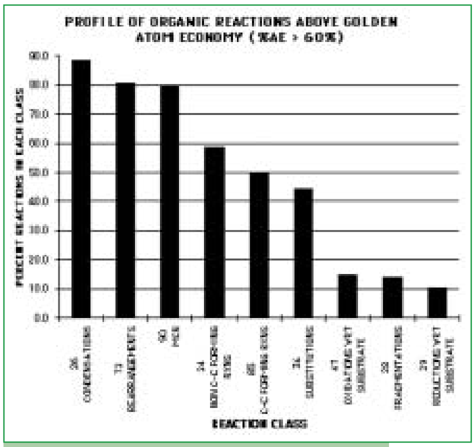


Figure 3. Fraction of each organic reaction class that meets the golden atom economy threshold.

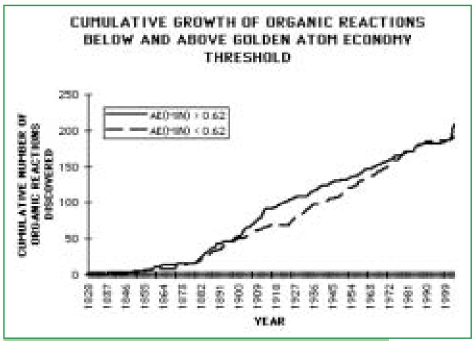


Figure 4. Cumulative growth of discovery of organic reactions below and above golden atom economy threshold since Wöhler's urea synthesis in 1828

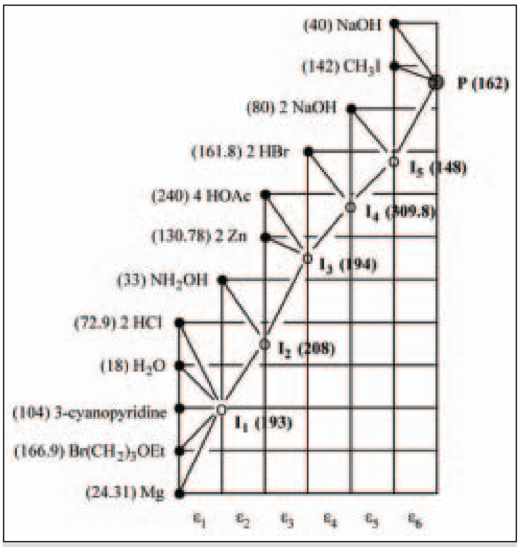


Figure 5. Synthesis tree for Scheme 2 to produce racemic nicotine

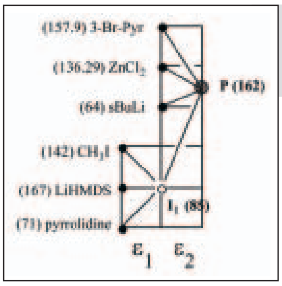
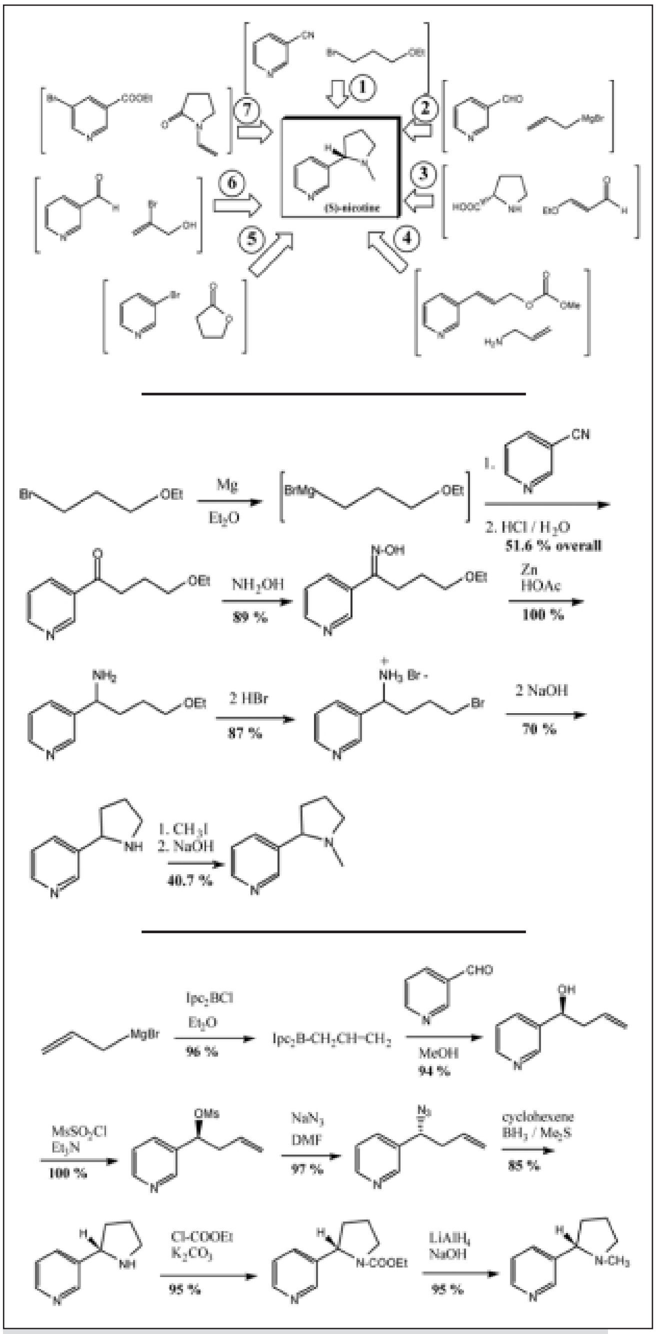


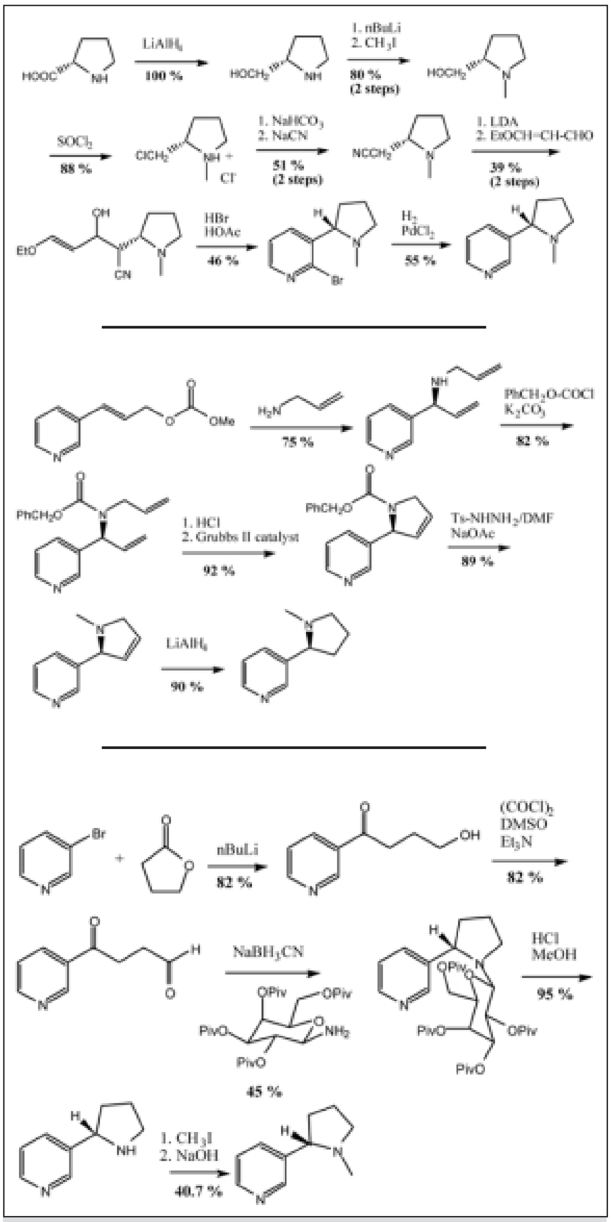
Figure 6. Synthesis tree for Scheme 9 to produce (+)(S)-nicotine

efficient plans to a given target and to focus attention on ones that have promising optimal RMEs. It is important to mention that the strength of the method is in the comparative sense rather than in the absolute sense. In simple language the new paradigm in synthesis design is to plan for a multi-component convergent strategy from the outset using benign starting materials and reagents in reactions that produce little or no byproducts and in high yield. The best goal to achieve in a synthesis plan is for each reaction in the plan to have an individual RME exceeding 62 percent. The criteria for good synthesis plans are therefore to:

- – Design convergent-reflexive plans from the outset that maximize the frequency of multi-component/tandem/domino/cascade reactions and minimize the frequency of redox reactions. Reflexive plans involve the production of a key source or "hub" intermediate which is then shunted to the production of other intermediates down the line which in turn are combined in a late-stage convergent step toward the target. A key strategy in deciphering such a hub intermediate is to look for the largest repeated structural motif in the target product. This is where exploitation of structural symmetries in the target molecule becomes important.
- – Design plans with as many MCRs as possible; overall convergence is maximized dramatically if a plan contains a late-stage MCR that involves components that are themselves terminal products of independent parallel linear sequences.
- – Design plans with large negative molecular weight first moments ($\mu 1$) (i.e., increase degree of building up from light starting materials via progressively heavier intermediates to final target product)
- – Minimize number of overall steps (M) and number of reaction stages (N)
- – Maximize number of parallel steps/branches
- – Minimize overall number of input materials needed (I)
- – Maximize number of input materials that end up in product in any given step (i.e., maximize atom economy (AE) for each step)
- – Maximize reaction yield for each step (ε)
- – Maximize overall kernel reaction mass efficiency (RME)
- – Maximize degree of convergence (δ)
- – Minimize degree of asymmetry (β)
- – Maximize relative rate of convergence (ρ_{rel})
- – Maximize fraction of total input energy directed toward target product
- – Minimize overall raw materials cost (RMC)



Scheme 1. Disconnection routes to nicotine
Scheme 2. Preparation of racemic nicotine
Scheme 3. (+)-S-nicotine prepared using chiral reagent (borane)

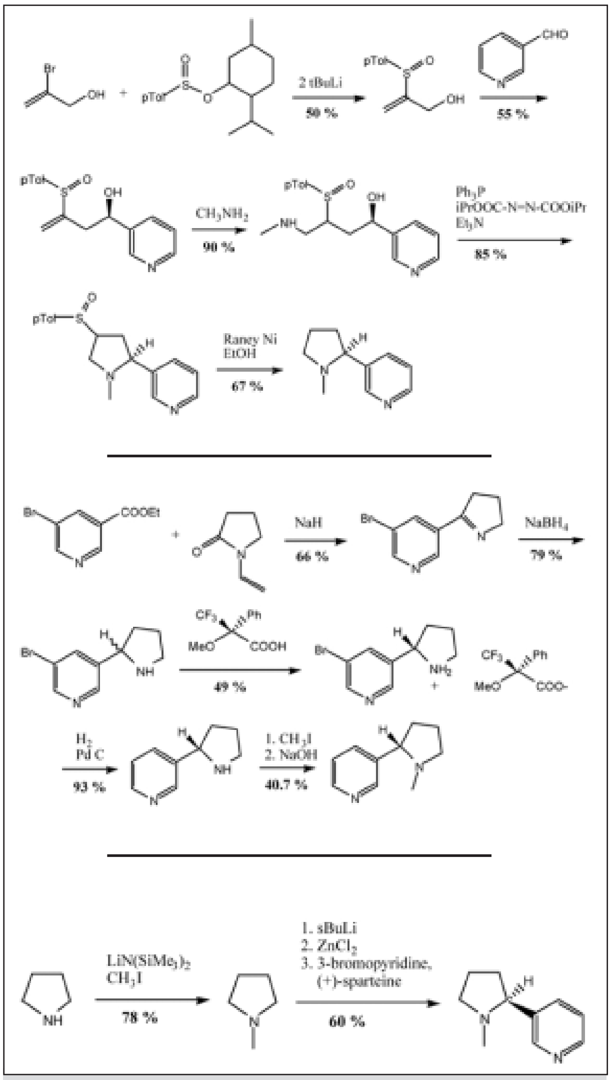


Scheme 4. (+)-S-nicotine prepared using chiral starting material (proline)
Scheme 5. (+)-S-nicotine prepared using chiral catalyst in step 1
Scheme 6. (+)-S-nicotine prepared using chiral auxiliary in step 3

The key point is that we desire all of these attributes in the *same* plan so that maximization and minimization of appropriate metrics occur synergistically!

AN ILLUSTRATED CASE: THE SYNTHESIS OF (+)(S)-NICOTINE

The above points may be convincingly illustrated by analyzing several literature plans for the synthesis of the well known alkaloid (+)(S)-nicotine (40 - 48). The molecule has two rings that are connected by a single bond. Strategies employed in the literature revolve around three main themes: (a) begin with a nicotinic acid derivative, extend the chain and cyclize to make the pyrrolidine ring and then methylate; (b) begin with a pyrrolidine ring precursor, extend the chain and cyclize to make the pyridine ring; and (c) begin with nicotinic acid and pyrrolidine derivatives (pyridine and pyrrole) and couple these stereoselectively. Scheme 1 summarizes various sets of starting materials for each route and Schemes 2 to 9 illustrate the specific synthesis plans. All plans are of the linear type. Scheme 2 shows the synthesis of the racemic product and Schemes 3 to 9 show the stereoselective synthesis of the natural enantiomer. Schemes 2, 3, 5, and 6 follow the first strategy, Schemes 4 and 7 follow the second strategy, and Schemes 8 and 9 follow the third. It should be noted that Scheme 8 involves a resolution step whereas that in Scheme 9 is a direct stereoselective coupling. The complete numerical analysis of all plans is summarized in Table 1. For illustrative purposes only Figures 5 and 6 show the synthesis trees corresponding to Schemes 2 and 9 respectively. It can be readily seen that Scheme 9, based on final product material recovered regardless of stereochemistry, has all the key attributes described for an optimal synthesis in the same plan; that is, it has the highest overall RME, the fewest number of input materials, the highest degree of building up in the synthesis, the highest degree of convergence, the fastest rate of convergence, and is the cheapest to produce on a per gram basis. These criteria suggest that this plan is the "greenest" on the basis of material efficiency and input material cost alone. The consideration of any other factors within the framework of green chemistry, most notably energy consumption, toxicity of materials, safety of materials and procedures particularly on scale-up, and waste treatment, may change this conclusion. However, it should be noted that plans with high material efficiency and which utilize cheap, readily available, and safe starting materials are strong contenders for overall "green" plans.



Scheme 7. (+)-S-nicotine prepared using chiral auxiliary in step 1
Scheme 8. (+)-S-nicotine prepared using classical resolution in step 2
Scheme 9. (+)-S-nicotine prepared using chiral reagent in step 2

	Plan 1[b] Scheme 2 (linear)	Plan 2 Scheme 3 (linear)	Plan 3 Scheme 4 (linear)	Plan 4 Scheme 5 (linear)	Plan 5 Scheme 6 (linear)	Plan 6 Scheme 7 (linear)	Plan 7 Scheme 8 (linear)	Plan 8 Scheme 9 (linear)
RME	0.0250	0.0857	0.00766	0.109	0.0185	0.0274	0.0268	0.115
E_m	38.62	10.67	129.50	8.21	53.10	35.48	36.33	7.70
AE	0.133	0.116	0.159	0.175	0.114	0.119	0.191	0.220
E_{mw}	6.49	7.64	5.29	4.71	7.77	7.42	4.23	3.56
$\varepsilon_{pseudo-overall}$ ($\varepsilon_{overall}$)	0.114	0.671	0.035	0.453	0.117	0.141	0.097	0.524
Number of reaction inputs, I	12	16	14	8	13	11	10	6
Number of reaction steps, M	6	7	7	5	5	6	5	2
Number of reaction stages, N	6	7	7	5	5	6	5	2
μ_1 (g per mole per reaction stage)	+ 27.10	+ 51.71	- 23.40	+ 33.33	+ 65.65	+ 43	+ 72.94	- 62.67
RMC[c] ($ CAD per gram)	13.65	12.31	68.13	26.75	42.82	202.97	922.10	4.98
P co-ordinate	(6, 9.7266)	(7, 13.287)	(7, 11.833)	(5, 5.742)	(5, 10.661)	(6, 8.835)	(5, 7.727)	(2, 3.375)
θ_p (degrees)	70.314	75.970	68.855	63.073	79.864	66.807	71.381	98.443
θ_{mcr} (degrees)	159.390	164.811	162.508	148.109	161.075	157.380	154.942	136.397
Degree of convergence δ	0.441	0.461	0.424	0.426	0.496	0.424	0.461	0.722
Pactual	1.621	1.898	1.69	1.148	2.132	1.472	1.545	1.690
Pl-mcr	5.5	7.5	6.5	3.5	6	5	4.5	2.5
Relative rate of convergence, r_{rel}	0.295	0.253	0.260	0.328	0.355	0.294	0.343	0.675
Degree of asymmetry, β	0.768	0.772	0.821	0.641	0.777	0.767	0.717	0.350

a) See reference 38 for meaning and determination of all parameters given in the table.
b) Refers to racemic product.
c) Raw material costs (RMC) based on prices of bulk starting materials from 2003 – 2004 Aldrich Chemical Catalogue in Canadian dollars.

Table 1a. Summary of reaction metrics analysis for synthesis plans of (+)(S)-nicotine shown in Schemes 2 to 9

RECOMMENDATIONS FOR REPORTING SYNTHESES OF TARGETS AS "GREEN"

When such green metrics and tree analyses are applied to the next new target molecule and are reported in the literature, we believe this will represent a breakthrough in convincing chemists of the power and scope of green chemistry principles. Green principles will be directly translated into tangible numbers that chemists can use with confidence in their synthesis planning strategies. Such a report will be a milestone on the order of Woodward's landmark paper on the total synthesis of quinine in 1945 (49) that launched the concept of total synthesis and showed that the synthesis of complex targets from smaller molecules was indeed feasible. What we advocate is the following standardization of reporting syntheses in the literature especially if a claim of some aspect of "green chemistry" is employed: (a) give a complete scheme with reagents and yields for each step using balanced chemical equations, (b) give an accompanying synthesis tree diagram, and (c) report overall kernel AE and RME, molecular weight first moment (degree of building up parameter), degree of convergence, degree of asymmetry, and overall yield of longest branch. What would be the obvious benefits of carrying out synthesis planning in this fashion? Planning syntheses with the aid of such numerical measures directs chemists' thinking in iterating toward rapid optimization of plans rather than relying on hit-and-miss strategies. It essentially increases the chemical intelligence of synthesis planners when conducting thorough surveys of literature plans and keeping abreast with and making use of the latest advances in new reaction transformations. Finally, it allows a planner to break free from the confines of traditional methods, constraints, or reactions in order to reach a given target structure in the least time, energy, and waste material produced, and ultimately cost. It needs to be emphasized that the ultimate driver for green chemistry usage is not environmental altruism but reducing cost in all its forms while producing high quality products safely and with benign life cycles from cradle to grave.

REFERENCES AND NOTES

1. G.H. Posner, *Chem. Rev.* **86**, 831 - 844 (1986).
2. Ugi, A. Dömling, W. Hörl, Endeavour (New Series) 18(3), 115 - 122 (1994).
3. R.W. Armstrong, A.P. Combs, P.A. Tempest, S.D. Brown, T.A. Keating, *Acc. Chem. Res.* **29**, 123 - 131 (1996).

JOHN ANDRAOS*, JULIA IZHAKOVA

***Corresponding author**
Department of Chemistry
4700 Keele Street, York University
Toronto, ON M3J 1P3, Canada

4. I. Ugi, A. Dömling, B. Gruber, M. Almstetter, *Croatica Chem. Acta* **70**, 631 – 647 (1997).
5. M. Lombardo, C. Trombini, Seminars in Organic Synthesis, Società Chimica Italiana: Milan, 1998, p. 7 - 32.
6. I. Ugi, *Proc. Estonian Acad. Sci. Chem.* **47**, 107 - 127 (1998).
7. A. Dömling, E. Herdtweck, I. Ugi, *Acta Chem. Scand.* **52**, 107 – 113 (1998).
8. I. Ugi, M. Almstetter, M. Bock, A. Dömling, B. Ebert, B. Gruber, C. Hanusch-Kompa, S. Heck, K. Kehagia-Drikos, K. Lorenz, S. Papathoma, R. Raditschnig, T. Schmid, B. Werner, A. von Zychlinkski, *Croatica Chem. Acta* **71**, 527 – 547 (1998).
9. L. Weber, I. Katrin, M. Almstetter, *Synlett*, 366 - 374 (1999).
10. I. Ugi, A. Dömling, B. Werner, *J. Heterocyclic Chem.* **37**, 647 - 658 (2000).
11. C.O. Kappe, *Acc. Chem. Res.* **33**, 879 - 888 (2000).
12. A. Dömling, *Curr. Opin. Chem. Biol.* **4**, 318 - 323 (2000).
13. A. Dömling, I. Ugi, *Angew. Chem. Int. Ed.* **39**, 3168 - 3210 (2000).
14. I. Ugi, *Pure Appl. Chem.* **73**, 187 - 191 (2001).
15. A. Dömling, *Curr. Opin. Chem. Biol.* **6**, 306 - 313 (2002).
16. L. Weber, *Drug Discovery Today* **7**, 143 - 147 (2002).
17. L. Weber, *Curr. Med. Chem.* **9**, 2085 - 2093 (2002).
18. R.V.A. Orru, M. de Greef, *Synthesis* 1471 - 1499 (2003).
19. A. von Wangelin, H. Neumann, D. Gördes, S. Klaus, D. Strübing, M. Beller, *Chem. Eur. J.* **9**, 4286 - 4294 (2003).
20. I. Ugi, B. Werner, A. Dömling, Molecules, 8, 53 - 66 (2003).
21. C. Simon, T. Constantieux, J. Rodriguez, *Eur. J. Org. Chem.* 4957 - 4980 (2004).
22. D.B. Ramachary, C.F. Barbas III, *Chem. Eur. J.* **10**, 5323 - 5331 (2004).
23. D.J. Ramon, M. Yus, *Angew. Chem. Int. Ed.* **44**, 1602 - 1634 (2005).
24. D. Tejedor, D. González-Cruz, A. Santos-Expósito, J.J. Marrero-Tellado, P. de Armas, F. García-Tellado, *Chem. Eur. J.* **11**, 3502 – 3510 (2005).
25. J. Zhu, H. Bienyamé, "Multicomponent Reactions", Wiley-VCH, 2005.
26. A. Dömling, *Chem. Rev.* **106**, 17 – 89 (2006).
27. P.T. Anastas, J.C. Warner, "Green Chemistry: Theory and Practice", Oxford University Press, 1998, p. 11.
28. B.M. Trost, *Science* **254**, 1471 – 1477 (1991).
29. B.M. Trost, *Acc. Chem. Res.* **35**, 695 – 705 (2002).
30. B.M. Trost, *Angew. Chem. Int. Ed.* **34**, 259 – 281 (1995).
31. R.A. Sheldon, *ChemTech* **24(3)**, 38 – 47 (1994).
32. R.A. Sheldon, *Chem. Ind.* (London) 903 – 906 (1992).
33. R.A. Sheldon, *Chem. Ind.* (London) 12 – 15 (1997).
34. R.A. Sheldon, *Pure Appl. Chem.* **72**, 1233 – 1246 (2001).
35. J. Andraos, *Org. Process Res. Dev.* **9**, 149 – 163 (2005).
36. J. Andraos, *Org. Process Res. Dev.* **9**, 404 – 431 (2005).
37. J. Andraos, *Org. Process Res. Dev.* **9**, 519 (2005).
38. J. Andraos, *Org. Process Res. Dev.* **10**, 212 – 240 (2006).
39. J. Andraos, M. Sayed, *J. Chem. Educ.*, in press.
40. F.B. LaForge, *J. Am. Chem. Soc.* **50**, 2477 – 2483 (1928).
41. F.X. Felpin, S. Girard, G. Vo-Thanh, R.J. Robins, R.J., J. Villieras, J. Lebreton, *J. Org. Chem.* **66**, 6305 – 6312 (2001).
42. C.G. Chavdarian, E.B. Sanders, R.L. Bassfield, *J. Org. Chem.* **47**, 1069 - 1073 (1982).
43. C. Welter, R.M. Moreno, S. Streiff, G. Helmchen, *Org. Biomol. Chem.* **3**, 3266 - 3268 (2005).
44. T.P. Loh, J.R. Zhou, X.R. Li, K.Y. Sim, *Tetrahedron Lett.* **40**, 7847 - 7850 (1999).
45. F. Marquez, A. Llebaria, A. Delgado, *Tetrahedron Asymmetry* **12**, 1625 - 1634 (2001).
46. P. Jacob III, *J. Org. Chem.* **47**, 4165 - 4167 (1982).
47. K.R. Campos, A. Klapers, J.H. Waldman, P.G. Dormer, C.Y. Chen, *J. Am. Chem. Soc.* **128**, 3538 - 3538 (2006).
48. R.B. Woodward, W. Doering, *J. Am. Chem. Soc.* **67**, 860 - 874 (1945).