Organic Process Research & Development **2005**, *9*, 149–163

Unification of Reaction Metrics for Green Chemistry: Applications to Reaction Analysis

John Andraos*

Department of Chemistry, York University, Ontario M3J 1P3 Canada

Abstract:

A formalism is presented which unifies key reaction metrics associated with "greenness" in chemical reactions with respect to raw materials usage. The fundamental basis of this treatment begins with balanced chemical reactions in which byproducts are identified. The primary or kernel metrics are reaction yield, scale of reaction, stoichiometric factor (SF), and Trost's atom economy (AE). The stoichiometric factor is a new metric that is defined to account for reactions run under nonstoichiometric conditions, that is, with one or more reagents in excess. A general relation for reaction mass efficiency (RME) is derived which shows that this metric is a composite of the aforementioned primary metrics and takes into account solvent usage in the reaction and postreaction phases (workup and purification). The Sheldon environmental impact factor *E* is treated at various levels of complexity according to what is constituted as waste and is shown to be related to RME by a simple inverse expression. A flowchart is presented which shows other simple relationships connecting all metrics. Raw material costs, optimum conditions for recycling or reclaiming catalysts and reaction and postreaction solvents, and the handling of reactions giving isomeric products are also assessed. General algorithms are proposed for determining kernel reaction metrics for linear and convergent sequences that can be used to compare the intrinsic, or best-case scenario, green performances of synthetic plans to a common target structure. All key relationships can be implemented in a spreadsheet format from which reaction histograms or "maps" can be generated. Individual reaction RME performances can be gauged, ranked, and decomposed according to AE, SF, and reaction yield kernel metrics. This allows for the easy identification of best and worst reactions in a process or sequence. Example applications of the present methodology include the following: (a) a comparative analysis of the synthesis of quinine by the classic Woodward−Rabe and the modern greener Stork methods; (b) the analysis of the industrial synthesis of sildenafil (Viagra) by a convergent strategy; and (c) the analysis of kinetic resolution of racemic alcohols by a successive oxidation and recycling reduction cycle.

1. Introduction

The need to invent efficient and environmentally friendly or "green" chemical reactions and processes is now a vigorous endeavour in organic chemistry. Coupled with this is the need to measure and rank the "green" performance of chemical reactions as quantitatively as possible. In this regard there has been considerable discussion in the literature concerning the quantification of "greenness" of chemical reactions using so-called green metrics. Among the most noted are Trost's atom economy[1] (AE), Sheldon's environmental impact factor[2] (*E*), and reaction mass efficiency[3] (RME). There have been recent studies[4] that reported the implementation of these reaction metrics to reaction sequences particularly those used in the pharmaceutical industry, since this chemical industry has been singled out as the one producing the most waste per gram of target product.[2b] These reports suggested that RME is a better and more useful metric than AE in gauging reaction performance. The main emphasis of the pharmaceutical industry's improved performance in reaction synthesis of targets was the demonstration of considerable waste reduction by recycling or eliminating reaction solvents and by reducing the number of isolations of intermediates along a given pathway by telescoping or concatenating one reaction into the next.[4]

Though the above metrics have been known and used for some time they have been presented as separate and unrelated quantities. Moreover there is still ongoing debate as to what metric is best to measure "greenness". Other competing metrics suggested by others include mass intensity (MI), carbon efficiency (CE), atom utilization (AU), environmental or elegance quotient (EQ), and mass efficiency.[5] The oldest reaction metric of course is percent yield with respect to the intended target product. In an effort to streamline and make useful the implementation of reaction metrics in "green" synthesis design, this report presents a detailed quantitative analysis of reaction metrics with respect to raw materials used in chemical reactions. Energy consumption, toxicities of materials, and hazards of processes are not included in this treatment.

The treatment begins with a simple analysis of a general balanced chemical equation to define kernel metrics and then increases in complexity as more variables are introduced to

---

* To whom correspondence should be addressed. E-mail: jandraos@yorku.ca. Fax: 1-416-736-5936.

(1) (a) Trost, B. M. *Science* **1991**, *254*, 1471. (b) Trost, B. M. *Acc. Chem. Res.* **2002**, *35*, 695. (c) Trost, B. M. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 259. (d) Sheldon, R. A. *Pure Appl. Chem.* **2001**, *72*, 1233.

(2) (a) Sheldon, R. A. *ChemTech* **1994**, *24* (3), 38. (b) Sheldon, R. A. *Chem. Ind. (London)* **1992**, 903. (c) Sheldon, R. A. *Chem. Ind. (London)* **1997**, 12.

(3) Curzons, A. D.; Constable, D. J. C.; Mortimer, D. N.; Cunningham, V. L. *Green Chem.* **2001**, *3*, 1.

(4) Anderson, N. G. *Practical Process Research and Development*; Academic Press: San Diego, 2000.

(5) (a) Constable, D. J. C.; Curzons, A. D.; Cunningham, V. L. *Green Chem.* **2002**, *4*, 521. (b) Constable, D. J. C.; Curzons, A. D.; Freitas dos Santos, L. M.; Geen, G. R.; Hannah, R. E.; Hayler, J. D.; Kitteringham, J.; McGuire, M. A.; Richardson, J. E.; Smith, P.; Webb, R. L.; Yu, M. *Green Chem.* **2001**, *3*, 7. (c) Eissen, M.; Hungerbühler, K.; Dirks, S.; Metzger, J. *Green Chem.* **2003**, *5*, G25.

10.1021/op049803n CCC: $30.25

account for excess reagents, catalysts, and solvent usage in the reaction and postreaction operations, namely, workup and purification. A general expression for reaction mass efficiency (RME) is derived and is shown to be inversely related to the overall Sheldon environmental impact factor, *E*. This facilitates the estimation of reaction efficiency and waste production by short cutting computations. Raw material costs, the inclusion of solvents and catalysts with and without recycling, and the handling of reactions giving isomeric products are also examined. The analysis of single reactions is then applied generally to sets of chemical reactions in the form of linear and convergent sequences so that general relations for overall RME and overall *E* are determined. All key relationships can be implemented in a spreadsheet format from which reaction histograms or "maps" can be generated. This allows for the easy identification of best and worst reactions in a process or sequence. Moreover, individual reaction RME performances can be gauged, ranked, and decomposed so as to identify which factors contribute to high and low RME values. At the level of kernel reaction metrics, the method is applied to the following: (a) the classic syntheses of quinine by the Woodward–Rabe and Stork methods and the industrial synthesis of sildenafil (Viagra), respectively, and (b) the analysis of kinetic resolution of racemic alcohols by a successive oxidation and recycling reduction cycle. This allows for convenient best case scenario comparisons between synthetic plans.

2. Definitions and Derivations

The analysis begins with a balanced chemical equation in which all byproducts are identified. Core or kernel reaction metrics can be classified under two groups: experimental parameters and calculated parameters. Under experimental parameters we have reaction yield, ϵ, and reaction scale, x. Under calculated parameters we have environmental impact factor *based on molecular weight* (E_{mw}), molecular weight of target product, p, and a new metric called stoichiometric factor (SF). The stoichiometric factor is introduced to handle reactions that are conducted under conditions where one or more of the reagents in a chemical reaction are used in excess, that is, under nonstoichiometric conditions. An SF equal to unity means that the reaction is operated under stoichiometric conditions; otherwise SF > 1. From these five key metrics and a balanced chemical equation, we now demonstrate how Trost's atom economy (AE), reaction mass efficiency (RME), environmental impact factor *based on mass* (E_{m}), and mass of waste, w, are derived under various levels of complexity depending on what is considered as waste. Interconnecting relationships between metrics are derived which show the dependence of the true Sheldon environmental impact factor (E-factor), E, on overall RME and conditions when it approaches the limiting values of E_{mw} and E_{m}. Figure 1 is a flowchart that illustrates the interconnections between all of these reaction metrics for any chemical reaction and Figure 2 summarizes key algebraic expressions under the assumption that solvents and catalysts are recycled.

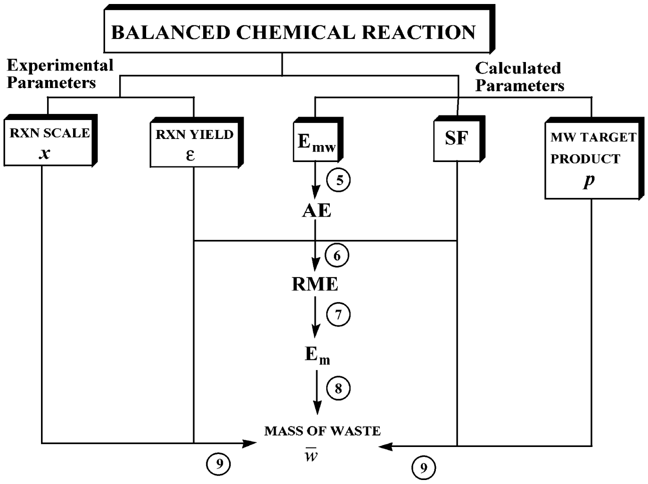


Figure 1. **Flowchart showing interconnections between kernel and peripheral reaction metrics for a single general chemical reaction under conditions where the solvent and any catalysts are recycled or reclaimed. Circled numbers refer to corresponding equation numbers in text.**

A

$$\boxed{RME = (\varepsilon)(AE)} \quad \boxed{E_m = \frac{1}{RME} - 1} \quad \boxed{\frac{AE}{RME} = \frac{1+E_m}{1+E_{mw}} = \frac{1}{\varepsilon}}$$

B

$$\boxed{RME = \frac{(\varepsilon)(AE)}{SF}} \quad \boxed{E_m = \frac{SF}{(\varepsilon)(AE)} - 1} \quad \boxed{\frac{AE}{RME} = \frac{1+E_m}{1+E_{mw}} = \frac{SF}{\varepsilon}}$$

Figure 2. **Summary of key kernel relationships among reaction metrics for a single general chemical reaction under conditions where the solvent and any catalysts are recycled. Group A refers to the set of expressions under stoichiometric conditions (SF = 1), and group B refers to the set of expressions under nonstoichiometric conditions (SF > 1). Symbols are defined in text.**

2a. Single Reaction Run under Stoichiometric Conditions Excluding Reaction and Postreaction Solvents. A general balanced chemical reaction is shown below where A represents the substrate of interest, B, the reagent, P, the target product, and Q, the waste byproduct.

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{Q} + \boxed{\mathrm{P}}$$

If a, b, p, and q represent the corresponding molecular weights of materials and the reaction is run under *stoichiometric conditions* with x mol each of A and B producing y mol each of P and Q, we may write an equation based on the law of conservation of mass given by

$$ax + bx = qy + py + w \tag{1}$$

where w is the mass of unreacted reactants A and B. The yield of the reaction is given by

$$\epsilon = \frac{y}{x} \tag{2}$$

In this first treatment, it is assumed that the reaction solvent, any catalysts used, and all solvents in postreaction operations are recycled or reclaimed, since their masses are not included in eq 1. In section 3 this constraint is removed. From the standard definitions of atom economy and environmental impact factor based on molecular weight, we have

$$AE = \frac{p}{a+b} = \frac{p}{q+p} \tag{3}$$

and

$$E_{\mathrm{mw}} = \frac{q}{p} \tag{4}$$

respectively. Note that $a + b = q + p$, since the chemical equation is balanced. In this notation AE is a dimensionless fraction between 0 and 1 and the units of E_{mw} are mass of waste of byproduct per mass of target product. It should be noted that both AE and E_{mw} are calculated using molecular weights of reactants and products, that is, units of grams per mole, so it is assumed that the scale of the reaction is 1 mol and that the reaction yield is 100%. It is straightforward to deduce the following connecting relationship between AE and E_{mw} from eqs 3 and 4.

$$AE = \frac{1}{1+E_{\mathrm{mw}}} \tag{5}$$

Equation 5 shows that atom economy and E_{mw} are inversely related. From the conservation of mass principle, this result is intuitively consistent with the idea that atom economy is a parameter measuring the fraction of reaction materials ending up in the target product, whereas E_{mw} is related to the fraction of reaction materials ending up as waste. Applying the definition of reaction mass efficiency (RME) as the ratio of mass of target product to the sum of masses of all reactants, we have

$$\mathrm{RME} = \frac{py}{ax+bx} = \frac{p\epsilon}{a+b} = \epsilon(AE) \tag{6}$$

The reaction mass efficiency is then the multiplicative product of the atom economy and reaction yield and is also a fraction between 0 and 1. We may define a corresponding environmental impact parameter based on mass (E_{m}) instead of molecular weight as follows (see Supporting Information for derivation):

$$E_{\mathrm{m}} = \frac{qy+w}{py} = \frac{1}{RME} - 1 \tag{7}$$

E_{m} is the ratio of mass of total waste per mass of target product and now takes into account the reaction yield, ϵ, in contrast to E_{mw}. In process development chemistry, this quantity is also known as the effluent load factor (ELF).[6] Here we see once again an inverse relationship between RME and E_{m} in the same sense as in eq 5 for AE and E_{mw}. The total mass of waste produced in the reaction, $\bar{w}$, corresponds to the sum of masses of the byproduct Q and unreacted starting materials A and B

$$\bar{w} = qy + w = pyE_{\mathrm{m}} = px\epsilon E_{\mathrm{m}} \tag{8}$$

Equation 8 may be written in a more convenient form (eq 9) that as will be seen later will be useful in handling linear sequences of reactions.

$$\bar{w} = xp\epsilon\left(\frac{1}{\mathrm{RME}} - 1\right) = xp\epsilon\left(\frac{1-\mathrm{RME}}{\epsilon(\mathrm{AE})}\right) = x\left(\frac{p}{\mathrm{AE}}\right)(1-\mathrm{RME}) \tag{9}$$

2b. Single Reaction Run under Nonstoichiometric Conditions Excluding Reaction and Postreaction Solvents. If the same chemical reaction is now run under nonstoichiometric conditions such that B is used in excess and A is the limiting reagent, then we proceed as follows. As before x mol of A are reacted but now with z mol of B such that $z > x$. We may define $\phi = z - x$ as the excess moles of B. The conservation of mass law becomes

$$ax + bz = qy + py + w \tag{10}$$

The definitions of reaction yield, AE, and E_{mw} remain the same as in eqs 2−4. However, the reaction mass efficiency now becomes

$$\mathrm{RME} = \frac{py}{ax+bz} = \frac{py}{ax+bx+b\phi} = \frac{p\epsilon}{a+b+b(\phi/x)} \tag{11}$$

If we define a stoichiometric factor, SF, as

$$\mathrm{SF} = 1 + \frac{b\phi}{ax+bx} \tag{12}$$

where $b\phi$ is the mass of excess B and $ax + bx$ is the stoichiometric mass of reactants, then eq 11 can be rewritten compactly as

$$\mathrm{RME} = \frac{\epsilon(\mathrm{AE})}{\mathrm{SF}} \tag{13}$$

Equation 13 is a more general expression for RME but still does not take into account catalyst and reaction solvent usage and other solvents used in postreaction operations such as workup and purification. Comparison with eq 6 shows that when at least one reagent is used in excess the reaction mass efficiency is further attenuated by the stoichiometric factor, since as can be seen from eq 12 SF has a magnitude of at least 1. Even at this level of complexity one observes that RME is subject to three competing factors. It is evident that reactions characterized by low reaction yields, low atom economy, nonstoichiometric operating conditions, or some combination of these parameters will have overall reduced reaction mass efficiencies. Conversely, reactions with high reaction mass efficiencies require high yields, high atom economies, stoichiometric operating conditions, or some combination of these parameters. The important point to recognize is that RME may be diminished by carrying out reactions under nonstoichiometric conditions even though reaction yields and/or atom economies may be high or by carrying out low atom economy reactions even though they may give high yields of product and/or are run under stoichiometric conditions. By analogy with the derivation of eq 7, E_{m} is given by

$$E_{\mathrm{m}} = \frac{\mathrm{SF}}{\epsilon(\mathrm{AE})} - 1 = \frac{1}{\mathrm{RME}} - 1 = \frac{\mathrm{SF}(1+E_{\mathrm{mw}})}{\epsilon} - 1 \tag{14}$$

which now takes into account reaction yield and excess

---

[6] Lee, S.; Robinson, G. *Process Development: fine chemicals from grams to kilograms*; Oxford University Press: Oxford, 1995; p 13.

reagents, and the total mass of waste is given by

$$\bar{w} = px\epsilon E_m = x\left(\frac{p}{AE}\right)(SF)(1 - RME) \qquad (15)$$

Note that eqs 13−15 reduce to eqs 6, 7, and 9 when SF = 1. Also, comparison of eqs 7 and 14 shows that the same inverse relationship between RME and E_m holds even under nonstoichiometric conditions.

3. Effect of Including Reaction and Postreaction Solvents and Catalysts and Effect of Recycling and Not Recycling Solvents and/or Catalysts

In the preceding analyses, the masses of all solvents and any catalysts used were not included and this is taken to mean that they are recycled or recovered. If they are not, then they will necessarily be included as part of the total mass of waste for the reaction. Hence the conservation of mass law and expressions for RME and E_m will be modified accordingly. Using the same chemical equation run under stoichiometric conditions, we have the conservation law

$$ax + bx + c + s + \omega = qy + py'' + w + (py - py'') + c + s + \omega \qquad (16)$$

where previous variable definitions remain the same and y'' is the number of moles of *actual* collected product after workup and purification, c is the mass of catalyst, s is the mass of reaction solvent, and ω is the mass of all solvents used in the workup and purification operations. The last parameter can be viewed as a "catch-all" mass parameter for all materials used in all postreaction operations, including solvents used in extractions and washes, acids and bases used in neutralization operations, and any other materials used in waste streams which may result in the overall synthesis of the desired target. The quantity $py - py''$ represents the mass of target product lost in workup and purification procedures. Following the same procedure as before, the generalized environmental impact factor based on mass is given by eq 17 assuming that the waste is composed of unreacted A and B, byproduct Q, catalyst, lost product during postreaction operations, and all solvents used (see Supporting Information for derivation):

$$E_m = \frac{qy + w + (py - py'') + c + s + \omega}{py''} = \frac{1}{\text{RME}} - 1 \qquad (17)$$

where, the generalized reaction mass efficiency is given by

$$\text{RME} = \langle\epsilon\rangle(\text{AE})\left[\frac{py''}{py'' + \langle\epsilon\rangle(\text{AE})(c + s + \omega)}\right] = \frac{1}{1 + E_m} \qquad (18a)$$

If any excess reagents are used then it can be shown, following the derivation of eq 13, that eq 18a is modified to eq 18b

$$\text{RME} = \frac{\langle\epsilon\rangle(\text{AE})}{\text{SF}}\left[\frac{py''}{py'' + \dfrac{\langle\epsilon\rangle(\text{AE})}{\text{SF}}(c + s + \omega)}\right] = \frac{1}{1 + E_m} \qquad (18b)$$

Table 1. **Reaction mass efficiency expressions for stoichiometric and nonstoichiometric reaction conditions for a variety of recycling scenarios**

Reaction solvent recycled	Reaction catalyst recycled	Post-reaction solvents recycled	RME[a] stoichiometric conditions	RME[a] non-stoichiometric conditions
Yes	Yes	Yes	$\langle\varepsilon\rangle(AE)$	$\dfrac{\langle\varepsilon\rangle(AE)}{SF}$
Yes	Yes	No	$\langle\varepsilon\rangle(AE)\left[\dfrac{py}{py+\omega\langle\varepsilon\rangle(AE)}\right]$	$\dfrac{\langle\varepsilon\rangle(AE)}{SF}\left[\dfrac{py}{py+\dfrac{\omega\langle\varepsilon\rangle(AE)}{SF}}\right]$
Yes	No	Yes	$\langle\varepsilon\rangle(AE)\left[\dfrac{py}{py+c\langle\varepsilon\rangle(AE)}\right]$	$\dfrac{\langle\varepsilon\rangle(AE)}{SF}\left[\dfrac{py}{py+\dfrac{c\langle\varepsilon\rangle(AE)}{SF}}\right]$
Yes	No	No	$\dfrac{\langle\varepsilon\rangle(AE)py}{py+(\Omega-s)\langle\varepsilon\rangle(AE)}$	$\dfrac{\langle\varepsilon\rangle(AE)py}{SF\left[py+\dfrac{(\Omega-s)\langle\varepsilon\rangle(AE)}{SF}\right]}$
No	Yes	Yes	$\langle\varepsilon\rangle(AE)\left[\dfrac{py}{py+s\langle\varepsilon\rangle(AE)}\right]$	$\dfrac{\langle\varepsilon\rangle(AE)}{SF}\left[\dfrac{py}{py+\dfrac{s\langle\varepsilon\rangle(AE)}{SF}}\right]$
No	Yes	No	$\dfrac{\langle\varepsilon\rangle(AE)py}{py+(\Omega-c)\langle\varepsilon\rangle(AE)}$	$\dfrac{\langle\varepsilon\rangle(AE)py}{SF\left[py+\dfrac{(\Omega-c)\langle\varepsilon\rangle(AE)}{SF}\right]}$
No	No	Yes	$\dfrac{\langle\varepsilon\rangle(AE)py}{py+(\Omega-\omega)\langle\varepsilon\rangle(AE)}$	$\dfrac{\langle\varepsilon\rangle(AE)py}{SF\left[py+\dfrac{(\Omega-\omega)\langle\varepsilon\rangle(AE)}{SF}\right]}$
No	No	No	$\langle\varepsilon\rangle(AE)\left[\dfrac{py}{py+\Omega\langle\varepsilon\rangle(AE)}\right]$	$\dfrac{\langle\varepsilon\rangle(AE)}{SF}\left[\dfrac{py}{py+\dfrac{\Omega\langle\varepsilon\rangle(AE)}{SF}}\right]$

[a] The masses of reaction solvent, catalyst, and postreaction solvents used in workup and purification operations are given by s, c, and ω, respectively; $\Omega = c + s + \omega$.

The symbol $\langle\epsilon\rangle$ represents the overall isolated reaction yield of target product P and is the multiplicative product of the kernel reaction yield from the reaction itself, workup extraction yield, and purification yield. It is observed that the definition of E_m in eq 17 more closely approximates the actual overall Sheldon impact factor, E, since all major waste contributors are taken into account and so E_m in equations 18a and 18b may be replaced by the symbol E, which is commonly used to represent this metric. *It is emphasized that the inverse relationship between RME and E_m or E is firmly established and arises directly from the law of conservation of mass for the entire process*. Also apparent from eqs 18a and 18b is that the atom economy metric is integral in the determination of the overall RME and, therefore, counters previous arguments[5a] that AE is less important than RME. In fact, eqs 18a and 18b show that the RME metric can be factored into three and four components, respectively, one of which is AE. Upon close inspection of the factor in square brackets in the above generalized RME expressions, one may already anticipate how any reclaiming or elimination of ancillary materials other than reagents in the reaction and postreaction processes will significantly improve the overall RME for a given reaction. This is discussed in more detail next.

It is clear that RME is an all-inclusive parameter that involves several factors all of which are important in its determination. Table 1 summarizes the resulting expressions for the RME under stoichiometric and nonstoichiometric conditions for a variety of recycling scenarios. It is observed that the base expressions for RME in each case are further attenuated by a factor that is less than unity and that the

Table 2. **Effect of recycling toluene and/or triethylamine on RME and E_m reaction metrics for the tosylation of benzyl alcohol described in the text**

solvent recovered (toluene)	catalyst recovered (triethylamine)	total mass waste (g)	E_m (g of waste/ g of product)	RME
yes	yes	9.11	0.39	0.721
yes	no	24.11	1.02	0.495
no	yes	509.11	21.57	0.044
no	no	524.11	22.21	0.043

masses of solvent or catalyst and ancillary materials denoted by the parameter ω appear in this attenuating factor when these materials are *not* recycled. The last entry in the nonstoichiometric conditions column is the most general expression for RME (corresponding to eq 18b) and would represent the minimum value of this metric, whereas the first entry in the stoichiometric conditions column represents its maximum value. One can deduce that the Sheldon environmental impact factor E for a general chemical reaction reduces to E_{mw} when the reaction yield, $\langle\epsilon\rangle$, approaches 100%, the reaction is run under stoichiometric conditions, $SF = 1$, any catalysts and solvents used are reclaimed, and all ancillary reaction and postreaction materials used are reclaimed or eliminated from the process.

The effect of recycling reaction solvents and/or catalysts can be illustrated numerically using a well documented example reaction[3,7] that has been used as an introductory teaching tool to explain reaction metrics analysis. The reaction involves treatment of 10.81 g (0.1 mol) of benzyl alcohol with 21.9 g (0.115 mol) of tosyl chloride in 500 g (5.427 mol) of toluene in the presence of 15 g (0.149 mol) of triethylamine to give 23.6 g (0.090 mol) of benzyl tosylate as product. In this example $E_{mw} = 0.139$, AE = 0.878, $\epsilon = 0.90$, and SF = 1.096. Table 2 summarizes the results of various recycling scenarios on RME and E_m according to expressions given in Table 1 under the assumption that in the postreaction operations all solvents are reclaimed and that no product losses occur. Since the solvent contributes the most mass in the reaction vessel, it is clear from the data in Table 2 that not recycling it has the most dramatic attenuating effect on the overall RME for the reaction. It is obvious from this analysis that reducing solvent usage as much as possible and/or reclaiming solvents is the most effective way of improving the reaction mass efficiency.

4. Effect of Recycling versus Retrieving Byproducts

In dealing with reactions that necessarily give byproducts in addition to the target product, one is faced with evaluating the merits of recycling versus retrieving versus disposal of the byproducts. The above formalism can be used to determine conditions when any one of these options may be advantageous. For simplicity it will be assumed that solvents and catalysts are recovered so that their masses will not be included in the analysis, and that reactions are carried out under stoichiometric conditions. The case of considering byproducts as waste, i.e., choosing the disposal option, has already been treated in Section 2a. If Q is retrieved, then it is not considered part of the waste. If, on the other hand, Q is recycled back to substrate A by some reaction given as

Table 3. **Summary of expressions for E_m, RME, and $\bar{w}$ for the three options dealing with byproduct Q[d]**

Condition	Total mass of waste, $\overline{w}$ (g)	E_m	RME
Q is treated as disposable waste[a]	$x\left(\frac{p}{AE}\right)\left(1-\varepsilon\left[(AE)-f_1\right]\right)$	$\frac{1}{\varepsilon(AE)}-1+f_1$	$\frac{\varepsilon(AE)}{1+f_1\varepsilon(AE)}$
Q is retrieved[b]	$x\left(\frac{p}{AE}\right)\left(1-\varepsilon\left[1-f_1\right]\right)$	$\left(\frac{1}{AE}\right)\left(\frac{1-\varepsilon}{\varepsilon}\right)+f_1$	$\frac{\varepsilon(AE)}{1+\varepsilon\left[(AE)(f_1+1)-1\right]}$
Q is recycled back to reagent A[c]	$x\left[\frac{p}{AE}+\varepsilon\left(C-a\varepsilon^*-p(1-f_2)\right)\right]$	$\frac{1}{\varepsilon(AE)}+\frac{C-a\varepsilon^*}{p}+f_2-1$	$\frac{p\varepsilon(AE)}{p+\varepsilon(AE)\left(C-a\varepsilon^*+pf_2\right)}$

[a] Conservation of mass law: $\bar{w} = ax + bx - py(1 - f_1)$. [b] Conservation of mass law: $\bar{w} = ax + bx - py(1 - f_1) - qy$. [c] Conservation of mass law: $\bar{w} = ax + bx + Cy - py(1 - f_2) - az$. [d] Symbol definitions: ϵ^* is reaction yield for recycling reaction equal to z/y; z is number of moles of Q* and A produced in the recycling reaction; C is the molecular weight of reagent C used in the recycling reaction; $f_1 = (c + s + \omega)/py$; $f_2 = f_1 + (c' + s' + \omega')/py$; c, s, and ω are the masses of catalyst, reaction solvent, and postreaction solvents for the initial product forming reaction; c', s', and ω' are the masses of catalyst, reaction solvent, and postreaction solvents for the recycling reaction.

$$\mathrm{Q} + \mathrm{C} \longrightarrow \mathrm{Q}^* + \boxed{\mathrm{A}}$$

where C is a new reagent and Q* is the byproduct of the recycling reaction, then its associated reaction metrics must be considered and evaluated in combination with the original reaction producing Q in the first place. A classic situation in which a recycling reaction may be considered is for redox reactions in which the substrate of interest is either oxidized or reduced by appropriate oxidizing or reducing agents. Such reactions necessarily result in the production of byproducts arising from the redox process and thus opportunities exist for regenerating these reagents, usually by another redox couple as the recycling reaction. Table 3 summarizes the corresponding expressions for E_m, RME, and $\bar{w}$ for the three options dealing with the byproduct Q assuming stoichiometric conditions throughout.

Under the simplified conditions of recovery of catalysts and all solvents used in the reaction and postreaction operations (f_1 and f_2 in Table 3 are set to zero; see definitions of these parameters in Table 3), it is possible to explore various recycling options with respect to the byproduct Q. It is obvious that retrieval of byproducts is always advantageous over committing them for disposal since $ax + bx - py - qy < ax + bx - py$ is always true. When comparing recycling byproducts versus maintaining them as waste, recycling is advantageous if $az > Cy$, that is, if the mass of recovered A is larger than the mass of C used in the recycling reaction. Alternatively the inequality is equivalent to $\epsilon^* > C/a$; that is, recycling is advantageous if the yield of the recycling reaction exceeds the ratio of the molecular weights of reagents C and A. Recycling becomes disadvantageous if $\epsilon^* < C/a$ or $az < Cy$. No advantage is imparted if $\epsilon^* =$

---

(7) (a) Cann, M. C.; Connelly, M. E. *Real World Cases in Green Chemistry*; American Chemical Society: Washington, DC, 2000. (b) Cann, M. C. The University of Scranton. http://academics.uofs.edu/faculty/CANNM1/organicmodule.html (accessed July 2004).

C/a or $az = Cy$. When comparing retrieval versus recycling of byproducts, retrieval is better than recycling if $qy + Cy > az$, that is, if the mass of recovered A is less than the sum of the masses of byproduct Q obtained from the first reaction and of reagent C used to recycle A. The inequality can be rewritten as $\epsilon^* < (q + C)/a$; that is, recycling Q is disadvantageous if the yield of the recycling reaction is less than the ratio of the sum of molecular weights of Q and C to that of reagent A. On the other hand recycling Q to A is better than retrieving it when $qy + Cy < az$ or $\epsilon^* > (q + C)/a$. If $qy + Cy = az$ or $\epsilon^* = (q + C)/a$, then there is no advantage. It can be shown that similar conclusions can be drawn for reactions run under nonstoichiometric conditions as well.

An important point to note is that inclusion of energy consumption in recycling and retrieving or recovery procedures needs to be considered. This can change dramatically the choice of possible outcomes. It is often the case that when energy demands and their associated costs are factored in, existing recycling options may be abandoned in favour of disposal. The viability of recycling, therefore, is strongly linked to the invention of energy efficient recycling reactions and processes. The value of the above analysis is that various options may be evaluated quantitatively at least within the constraint of raw materials consumption.

5. Raw Materials Costs

When considering costs of raw materials, excluding solvents and catalysts, for a general reaction, the scale of the reaction, x, needs to be first determined for a specified mass of target product, M. This is given by $x = M/p\epsilon$, where p and ϵ are the molecular weight of target product and reaction yield, respectively. For a general reaction following the form of that given in section 2a, if the costs per unit mass of reagents A and B are $\$_A$ and $\$_B$, then the required raw material cost (RMC) to produce a mass M of product P will be given by

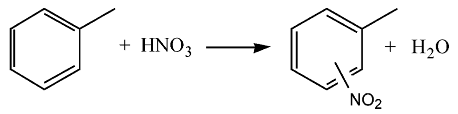
$$\text{RMC} = \frac{M(a\$_A + b\$_B)}{p\epsilon} = \frac{\text{RMC}_{100}}{\epsilon} \tag{19}$$

under stoichiometric operating conditions and by

$$\text{RMC} = \frac{M\left(a\$_A + b\$_B\left[1 + (\text{SF} - 1)\left(1 + \frac{a}{b}\right)\right]\right)}{p\epsilon} = \frac{\text{RMC}_{100}}{\epsilon} \tag{20}$$

under nonstoichiometric operating conditions, where a and b are the molecular weights of A and B, SF is the stoichiometric factor assuming that reagent B is used in excess, and RMC_{100} is the raw material cost at 100% reaction yield. Note that eq 20 reduces to eq 19 when SF = 1. The costs of catalysts, solvents, and other ancillary materials can be incorporated directly in each case by adding these to the RMC quantities for reagents obtained above. Other costs that can also be directly added include labor, energy, equipment (capital and/or maintenance), waste stream treatment, and opportunity costs. A referee has noted that, at least within the pharmaceutical industry, the costs per unit of raw materials often decrease as larger amounts of raw materials are purchased.[8]

Scheme 1



MW	92	63	137	18
moles	x	x	y	y
mass	$92x$	$63x$	$137y$	$18y$

6. Special Reactions

A number of special cases for reactions are now considered in the context of raw materials usage under the assumption that all catalysts and solvents are recycled or reclaimed, that is, within the determination of kernel reaction metrics.

6a. Case of Reactions Producing Several Isomeric Products. When dealing with reactions that produce several isomeric products, the analysis of kernel reaction mass efficiency for such a reaction is given generally by eqs 21−23.

$$\text{RME}_{\text{overall}} = \frac{\epsilon_{\text{overall}}(\text{AE})}{\text{SF}} = \sum_{j=1}^{r}(\text{RME})_j \tag{21}$$

$$(\text{RME})_j = f_j(\text{RME})_{\text{overall}} \tag{22}$$

$$\epsilon_j = f_j\epsilon_{\text{overall}} \tag{23}$$

where r is the number of isomeric products and f_j is the fraction of the jth isomeric product. A simple example is the nitration of toluene which produces a mixture of ortho-, meta-, and para-substituted nitrotoluenes in the general ratio $o/m/p$. Scheme 1 illustrates such a reaction operated under stoichiometric conditions (SF = 1). The atom economy is 0.884, and if the overall yield of nitrotoluenes is $\epsilon_{\text{overall}} = y/x$, then the overall RME to produce all the nitrotoluenes is $0.884(y/x)$. The individual kernel reaction mass efficiencies for each isomeric product are given by

$$(\text{RME})_{\text{ortho}} = 0.884\left(\frac{y}{x}\right)\left(\frac{o}{o + m + p}\right) \tag{24a}$$

$$(\text{RME})_{\text{meta}} = 0.884\left(\frac{y}{x}\right)\left(\frac{m}{o + m + p}\right) \tag{24b}$$

$$(\text{RME})_{\text{para}} = 0.884\left(\frac{y}{x}\right)\left(\frac{p}{o + m + p}\right) \tag{24c}$$

For example, if p-nitrotoluene was the desired product from such a reaction, then the ortho and meta isomers are considered as waste byproducts and hence the kernel RME for the process is given by eq 24c.

6b. Case of Reactions Producing Two Stereoisomeric Products. Kernel reaction mass efficiencies for reactions which produce pairs of stereoisomers may be found analogously according to eqs 25−27

---

(8) Anderson, N. G. *Org. Process Res. Dev.* **2004**, *8*, 260.

$$(\mathrm{RME})_{\mathrm{overall}} = \sum_{j=1}^{2} (\mathrm{RME})_j \tag{25}$$

$$(\mathrm{RME})_j = \begin{cases} \left(\dfrac{\epsilon_{\mathrm{overall}}(\mathrm{AE})}{\mathrm{SF}}\right)\left(\dfrac{1+\mathrm{ee}}{2}\right) \\ \left(\dfrac{\epsilon_{\mathrm{overall}}(\mathrm{AE})}{\mathrm{SF}}\right)\left(\dfrac{1-\mathrm{ee}}{2}\right) \end{cases} = \begin{cases} (\mathrm{RME})_{\mathrm{overall}}\left(\dfrac{1+\mathrm{ee}}{2}\right) \\ (\mathrm{RME})_{\mathrm{overall}}\left(\dfrac{1-\mathrm{ee}}{2}\right) \end{cases} \tag{26}$$

$$\epsilon_j = \begin{cases} \epsilon_{\mathrm{overall}}\left(\dfrac{1+\mathrm{ee}}{2}\right) \\ \epsilon_{\mathrm{overall}}\left(\dfrac{1-\mathrm{ee}}{2}\right) \end{cases} \tag{27}$$

where ee represents enantiomeric excess with respect to the R stereoisomer for enantiomerically related products.

6c. Case of Reactions Producing Two Stereoisomeric Products Involving Dynamic Kinetic Resolution. Scheme 2 illustrates a general stoichiometric reaction of stereochemically related substrates S1 and S2 in an initial ratio a/b that produce stereochemically related products P1 and P2 whose final ratio is given by c/d. The respective initial masses of S1 and S2 are $(a/(a+b))xs$ and $(b/(a+b))xs$, and the respective final masses P1 and P2 are $(c/(c+d))yp$ and $(d/(c+d))yp$. Supposing that the target product is P1, then the yield of this product is given by

$$\epsilon_{\mathrm{P1}} = \left(\frac{c}{c+d}\right)\left(\frac{y}{x}\right) = \left(\frac{c}{c+d}\right)\epsilon_{\mathrm{overall}} \tag{28}$$

and its corresponding kernel reaction mass efficiency is then given by

$$(\mathrm{RME})_{\mathrm{P1}} = \epsilon_{\mathrm{P1}}(\mathrm{AE}) = \epsilon_{\mathrm{P1}}\left(\frac{p}{s+r}\right) \tag{29}$$

The final product ratio c/d is of course governed by the efficiency of dynamic resolution which depends on the relative magnitudes of rate constants for the equilibration step between S1 and S2 and those for the product forming steps. A quantitative expression for the efficiency of dynamic kinetic resolution has recently been proposed, and its determination from experimentally determined substrate and product ratios has been investigated in the context of Curtin−Hammett and anti- Curtin−Hammett conditions.[9]

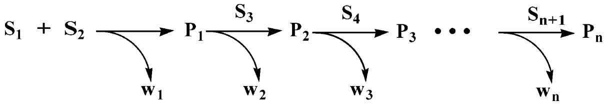
7. Application to Linear Sequences

We now apply the reaction metrics analysis discussed previously for individual reactions to linear sequences composed of consecutive reactions as depicted in Scheme 3. Linear sequences represent the majority of synthetic pathways to important target products.

Scheme 2

$$\begin{array}{lccccccc} & \mathbf{S1/S2} & + & \mathbf{R} & \longrightarrow & \mathbf{P1/P2} & + & \mathbf{Q} \\ \mathrm{MW} & s & & r & & p & & q \\ \mathrm{moles} & x & & x & & y & & y \\ \mathrm{mass} & sx & & rx & & py & & qy \end{array}$$

Scheme 3



In step j we have the following parameters: w_j, mass of waste of unreacted starting materials and byproducts); c_j, mass of catalyst; s_j, mass of solvent; p_j, molecular weight of intermediate product; and ϵ_j, fractional yield. The scale of the reaction sequence, x, is defined by the number of moles of limiting reagents used in the first step of the sequence. An implicit assumption is that each intermediate product along the pathway is committed as a substrate in the next consecutive step. Such an assumption consequently results in the overall yield to be the multiplicative product of individual reaction yields for each step. It should be noted that this is the *only* condition which validates multiplying yields of successive reactions in a linear sequence. Generally, the overall Sheldon environmental impact factor, E_{overall}, for a linear sequence is given by

$$E_{\mathrm{overall}} = \frac{\sum_{j=1}^{n} \bar{w}_j}{p_n x(\prod_{j=1}^{n} \epsilon_j)} = \frac{\sum_{j=1}^{n} (w_j + q_j y_j + c_j + s_j + \omega_j)}{p_n x(\prod_{j=1}^{n} \epsilon_j)} \tag{30}$$

where x is the scale of the reaction defined by the first step in the sequence, $p_n x$ is the theoretical mass of final product, $\prod_{j=1}^{n} \epsilon_j$ is the overall reaction yield for the n steps, and w_j, $q_j y_j$, c_j, s_j, and ω_j are the masses of unreacted reagents, byproduct, catalyst, reaction solvent, and postreaction solvents used in the jth step, respectively. From eq 30 the overall reaction mass efficiency may be determined directly by eq 31 (see Supporting Information for derivation).

$$\mathrm{RME}_{\mathrm{overall}} = \frac{1}{1 + E_{\mathrm{overall}}} \tag{31}$$

The mass of waste produced in the first step and succeeding steps are given by eqs 32 and 33, respectively.

$$\bar{w}_1 = x\left(\frac{p_1}{(\mathrm{AE})_1}\right)[(\mathrm{SF})_1 - \epsilon_1(\mathrm{AE})_1] + s_1 + c_1 + \omega_1, \text{ for } j = 1 \tag{32}$$

$$\bar{w}_j = x\left(\frac{p_j}{(\mathrm{AE})_j}\right)(\prod_{k=1}^{j-1} \epsilon_k)[(\mathrm{SF})_j - \epsilon_j(\mathrm{AE})_j] + s_j + c_j + \omega_j, \text{ for } j > 1 \tag{33}$$

Alternatively, the mass of waste produced in each step in the sequence can be determined from eq 34.

$$\bar{w}_j = E_j\, p_j\, x(\prod_{k=1}^{j} \epsilon_k) \tag{34}$$

The stoichiometric factor for step j may be generalized to eq 35 if more than one reagent is used in excess in that step.

---

(9) Andraos, J. *J. Phys. Chem. A* **2003**, *107*, 2374.

$$(SF)_j = 1 + \frac{\sum \text{masses excess reagents}}{\sum \text{masses reagents with no excesses}} = 1 + \frac{\sum_{k=1} b_k \phi_k}{x(a + b + ...)} = 1 + \frac{(\mathrm{AE})_j \sum_{k=1} b_k \phi_k}{x p_j} \tag{35}$$

Other useful relationships relating reaction metrics for the overall process to individual steps are given below in eqs 36−39.

$$E_{\mathrm{mw}}^{\mathrm{overall}} = \sum_{j=1}^{n} E_{\mathrm{mw}}^{j} \left(\frac{p_j}{p_n}\right) \tag{36}$$

$$(\mathrm{AE})_{\mathrm{overall}} = \frac{1}{1 + E_{\mathrm{mw}}^{\mathrm{overall}}} = \frac{1}{1 + \sum_{j=1}^{n} \left(\frac{p_j}{p_n}\right)\left(\frac{1}{(\mathrm{AE})_j} - 1\right)} \tag{37}$$

$$E_{\mathrm{overall}} = \sum_{j=1}^{n} E_j \left(\frac{p_j}{p_n}\right)\left(\frac{1}{\epsilon_n \epsilon_{n-1} ... \epsilon_{j+1}}\right) \tag{38}$$

$$(\mathrm{RME})_{\mathrm{overall}} = \frac{1}{1 + E_{\mathrm{overall}}} = \frac{1}{1 + \sum_{j=1}^{n} \left(\frac{1}{(\mathrm{RME})_j} - 1\right)\left(\frac{p_j}{p_n}\right)\left(\frac{1}{\epsilon_n \epsilon_{n-1} ... \epsilon_{j+1}}\right)} \tag{39}$$

If a mass, M, of target final product is intended using the sequence, then the overall raw materials cost for the process excluding solvents and catalysts is given by

$$\begin{aligned}(\mathrm{RMC})_{\mathrm{overall}} &= \sum_{j=1}^{n} \frac{(\mathrm{RMC})_{j,100}}{\epsilon_j \epsilon_{j+1} ... \epsilon_n} \\ &= \frac{(\mathrm{RMC})_{1,100}}{\epsilon_1 \epsilon_2 ... \epsilon_n} + \frac{(\mathrm{RMC})_{2,100}}{\epsilon_2 \epsilon_3 ... \epsilon_n} + ... + \frac{(\mathrm{RMC})_{n,100}}{\epsilon_n}\end{aligned} \tag{40}$$

where the first term in the sum corresponds to the first step, etc. The raw materials costs at 100% yield for the first and successive steps are given by

$$(\mathrm{RMC})_{1,100} = \left(\frac{M}{p_n}\right)\left[s_1 \$_{S_1} + s_2 \$_{S_2}\left\{1 + ((\mathrm{SF})_1 - 1)\left(1 + \frac{s_1}{s_2}\right)\right\}\right] \tag{41}$$

and

$$(\mathrm{RMC})_{j,100} = \left(\frac{M}{p_n}\right)\left[s_{j+1} \$_{S_{j+1}}\left\{1 + ((\mathrm{SF})_j - 1)\left(1 + \frac{p_{j-1}}{s_{j+1}}\right)\right\}\right], j > 1 \tag{42}$$

where the parameters have the same meaning as given in eqs 19 and 20 and the symbolism corresponds to that given in Scheme 3; i.e., $\$_{S_{j+1}}$ is the cost per unit mass of reagent S_{j+1} in step j, s_{j+1} is the molecular weight of reagent S_{j+1} in step j, and p_{j-1} is the molecular weight of product obtained in the step preceding the jth step.

A simplified algorithm for determining the overall reaction mass efficiency and cumulative mass of waste for any linear sequence under the assumption of reclaimed reaction and postreaction solvents can be represented by the following series of steps:

1. Balance chemical equations accounting for all byproducts in each step.
2. Determine E_{mw}^{j} for each step.
3. Determine $(\mathrm{AE})_j$ for each step using $(\mathrm{AE})_j = 1/(1 + E_{\mathrm{mw}}^{j})$.
4. Determine $E_{\mathrm{mw}}^{\mathrm{overall}} = \sum$ MW byproducts/MW target product.
5. Determine $(\mathrm{AE})_{\mathrm{overall}} = 1/(1 + E_{\mathrm{mw}}^{\mathrm{overall}})$.
6. Determine $(\mathrm{SF})_j = 1 + [(\mathrm{AE})_j \sum_{k=1} b_k \phi_k / x_j p_j]$.
7. Determine $(\mathrm{RME})_j = \epsilon_j (\mathrm{AE})_j / (\mathrm{SF})_j$.
8. Determine $E_{\mathrm{m},j} = [1/(\mathrm{RME})_j] - 1$.
9. Determine mass of waste in step j using
   $\bar{w}_j = E_{m,j} p_j x (\prod_{k=1}^{j-1} \epsilon_k)$ or
   $\bar{w}_1 = x(p_1/(\mathrm{AE})_1)[(\mathrm{SF})_1 - \epsilon_1 (\mathrm{AE})_1]$, for $j = 1$
   $\bar{w}_j = x(p_j/(\mathrm{AE})_j)(\prod_{k=1}^{j-1} \epsilon_k)[(\mathrm{SF})_j - \epsilon_j (\mathrm{AE})_j]$, for $j > 1$
10. Determine $\sum_{j=1}^{n} \bar{w}_j$.
11. Determine $E_{\mathrm{m}}^{\mathrm{overall}} = [\sum_{j=1}^{n} \bar{w}_j]/[p_n x (\prod_{k=1}^{j-1} \epsilon_k)]$.
12. Determine $(\mathrm{RME})_{\mathrm{overall}} = 1/(1 + E_{\mathrm{m}}^{\mathrm{overall}})$.

If solvents and catalysts are not recovered in any of the steps, then expressions appearing in item 9 are replaced by eqs 32 and 33 accordingly. It is clear that solvent recovery necessarily maximizes overall RME and minimizes overall E_{m}. It is important to note that the overall reaction mass efficiency is *not* equal to the multiplicative product of the overall atom economy and overall reaction yield. This will become clear in the following numerical demonstration.

The value of analyzing synthetic plans by this quantitative approach is that they can be compared directly under best case scenario conditions, that is, if all solvents and catalysts are recovered or if indeed a process involves solventless steps. This allows for the direct comparison of the intrinsic *chemical* performances of each reaction based on reaction yield, atom economy, reaction mass efficiency, and whether reactions are run with excess reagents or not. If an analysis for a given synthesis plan is not favourable at the kernel metrics stage, then as was shown by entries in Table 1 it is certain that things can only get worse when solvents and other ancillary materials in the reaction and postreaction processes are taken into account. It is imperative therefore to choose a synthesis plan for a given target that at the outset has a combination of high reaction yields, high atom economies, and SF values close to unity for each step. As an illustrative example of the analysis of kernel metrics, the experimental results of the Woodward−Rabe[10] and Stork[11] syntheses of quinine are analyzed and compared according to the above algorithm and the results are summarized in Tables 4 and 5. This was chosen as a test of the method since the total synthesis of quinine is widely regarded as the

---

[10] (a) Woodward, R. B.; Doering, W. E. *J. Am. Chem. Soc.* **1945**, *67*, 860. (b) Rabe, P. *Chem. Ber.* **1911**, *44*, 2088. (c) Rabe, P.; Huntenburg, W.; Schultze, A.; Volger, G. *Chem. Ber.* **1931**, *64*, 2487.

[11] Stork, G.; Niu, D.; Fujimoto, A.; Koft, E. R.; Balkovec, J. M.; Tata, J. R.; Dake, G. R. *J. Am. Chem. Soc.* **2001**, *123*, 3239.

Table 4. **Summary of kernel reaction metrics analysis of Woodward−Rabe total synthesis of (±)quinine determined by the algorithm described in text under the assumptions of recoverable reaction solvents and catalysts, solvents used in workup and purification procedures, and each intermediate product committed as a substrate in the next consecutive step in the sequence**[a]

step no.	reaction type	MW, p_j	ϵ	SF	E_{mw}	AE	RME	E_{m}	mass waste (g)[b]
1	reduction + diazotization + substitution	122	0.64 [c]	1.30	9.80	0.093	0.05	20.93	1081.83
2	condensation	237	0.94	1.03	0.08	0.93	0.85	0.18	16.82
3	elimination of EtOH	145	0.64	1	0.64	0.61	0.39	1.56	57.71
4	C−C bond forming	242	0.61	1.01	0.08	0.93	0.56	0.78	29.36
5	elimination of piperidine	159	0.65	9.73	1.21	0.45	0.03	32.14	516.45
6	reduction	163	0.67	3.41	0	1	0.20	4.09	45.14
7	N-acylation	205	0.66	1.08	0.30	0.77	0.47	1.13	10.31
8	reduction	211	0.5	>1	0	1	<0.50	>1	4.71
9	oxidation	227	0.52	4.93	0.89	0.53	0.06	16.89	44.54
10	elimination of water + oximation	284	0.71[d]	8.3	0.25	0.8	0.07	13.61	31.89
11	reduction	270	1	1	0.06	0.94	0.94	0.06	0.14
12	methylation	440	0.91[d]	4.45	0.59	0.63	0.13	6.76	22.34
13	Hofmann elimination + urea formation	226	0.38[d]	3.2	1.86	0.35	0.04	23.06	14.87
14	benzamide formation	315	0.96	29.4	0.56	0.64	0.02	46.85	40.42
15	condensation	500	0.63	1.53	0.23	0.81	0.33	2	1.72
16	double hydrolysis + decarboxylation	324	0.79	11.2	0.67	0.6	0.04	22.63	9.99
17	N-bromination	402.9	0.31	1.13	0.41	0.71	0.19	4.13	0.70
18	cyclization + debromination	322	0.46	3	0.47	0.68	0.10	8.59	0.54
19	reduction	324	1[e]	1.3	0.54	0.65	<0.50	>1	0.06
overall			<0.029%		11.0	0.08	<0.0032%	>31 020	1929.5

[a] Molecular weight quinine product, $p = 324$. [b] Scale of reaction, $x = 0.662$ mol. [c] Combined yield for three steps. [d] Combined yield for two steps. [e] Not reported; assumed to be 100% yield.

Table 5. **Summary of kernel reaction metrics analysis of Stork total synthesis of (−)quinine determined by the algorithm described in text under the assumptions of recoverable reaction solvents and catalysts, solvents used in workup and purification procedures, and each intermediate product committed as a substrate in the next consecutive step in the sequence**[a]

step no.	reaction type	MW, p_j	ϵ	SF	E_{mw}	AE	RME	E_{m}	mass waste (g)[b]
1	ring opening amidation	185	1	1.39	0	1	0.72	0.39	1.52
2	protection of alcohol	299	0.79[c]	1.2	0.54	0.65	0.43	1.34	6.63
3	C-alkylation with protecting group	581	0.79	1.07	0.52	0.66	0.49	1.05	8.01
4	deprotection	467	0.93	13.19	0.35	0.74	0.05	18.17	103.40
5	cyclization + deamination	394	0.97	1	0.19	0.84	0.81	0.23	1.06
6	reduction	396	1	7.89	0.41	0.71	0.09	10.11	47.35
7	ring opening + Wittig	424	0.75[c]	1.66	1.22	0.45	0.20	3.92	14.73
8	azide formation	449	0.95	1.15	1.56	0.39	0.32	2.10	7.96
9	hydration of vinyl ether	435	0.78	3.85	0.08	0.93	0.19	4.31	12.31
10	C-alkylation	608	0.72	12.2	0.20	0.83	0.05	19.41	55.84
11	Swern oxidation	606	0.85	>1.86	0.54	0.65	<0.30	>2.37	5.77
12	reductive cyclization	562	0.81	1.03	0.54	0.65	0.51	0.96	1.75
13	reduction	564	0.91	10.87	0.04	0.96	0.08	11.44	19.13
14	deprotection	326	0.95	1	0.79	0.56	0.53	0.88	0.81
15	mesylation	404	1	1.31	0.28	0.78	0.60	0.68	0.77
16	cyclization + elimination of MsOH	308	0.68[c]	1	0.32	0.76	0.52	0.93	0.55
17	oxidation	324	0.78	>13.3	0.25	0.8	<0.05	20.31	9.83
overall			7.11%		10.6	0.086	<0.16%	>614.6	297.42

[a] Molecular weight quinine product, $p = 324$. [b] Scale of reaction, $x = 0.021$ mol. [c] Combined yield for two steps.

genesis of the art of organic synthesis as a means to construct complex target structures from simple readily available materials using the growing library of organic reactions discovered over time. Figures 3−5 depict various comparative graphs or “reaction maps” showing the “green” performance profiles for each synthetic route according to the kernel metrics: reaction yield, atom economy, and stoichiometric and reaction mass efficiency under the assumption

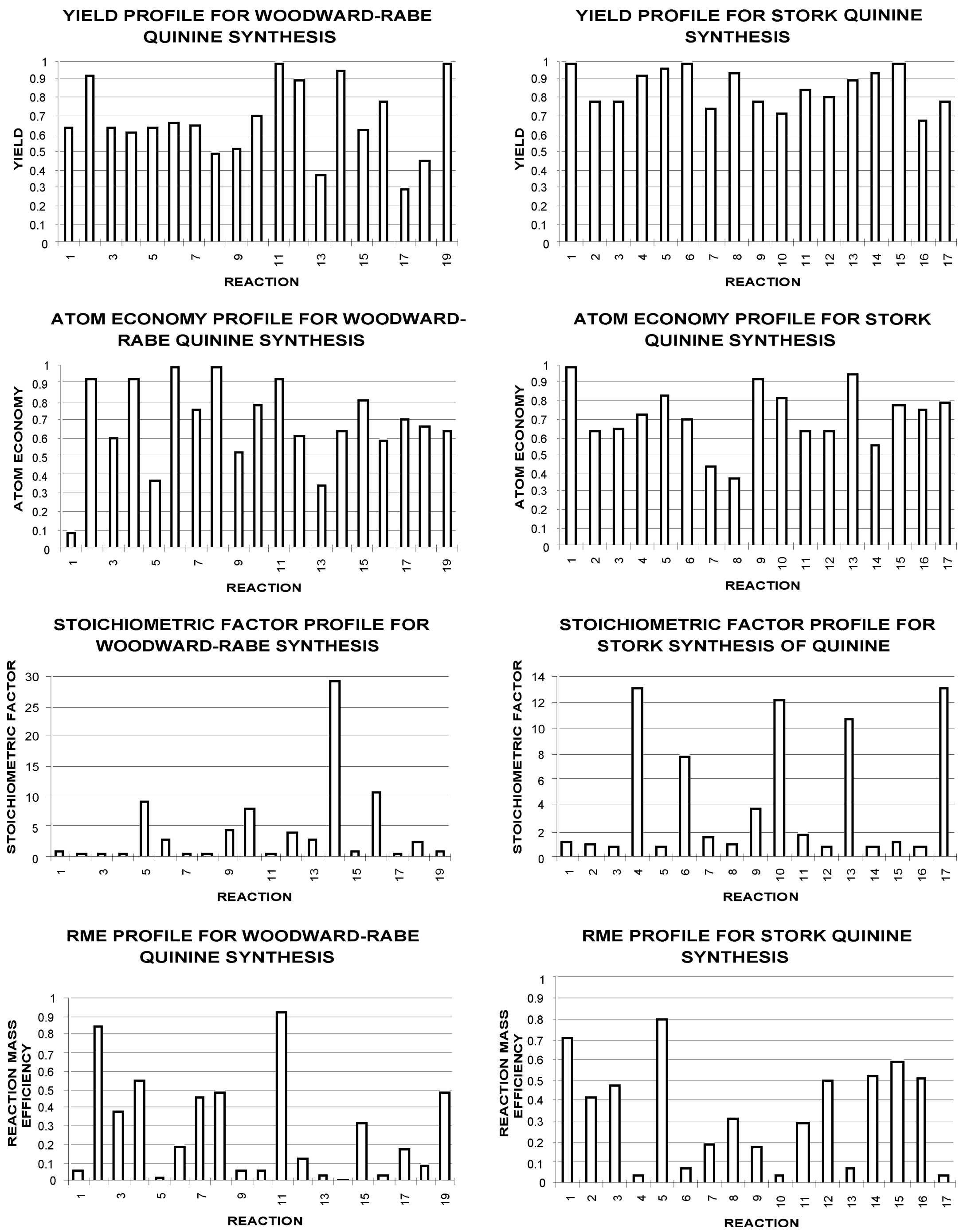


Figure 3. **Yield, atom economy, stoichiometric factor, and reaction mass efficiency profiles or reaction maps for Woodward−Rabe and Stork syntheses of quinine. Reaction mass efficiencies are determined using eq 13.**

that all solvents and catalysts are recoverable. Reaction schemes for the two methods showing all products are given in the Supporting Information.

Table 6 compares the two synthetic strategies to quinine according to overall performance metrics and reaction classes employed. The primary difference between the two synthetic strategies is that the Stork method stereoselectively targets (−)-quinine, whereas the Woodward−Rabe one affords the racemic product. In the present numerical illustration of the algorithm, the focus is on the determination of the absolute amount of material collected in each case irrespective of the stereochemistry. It is interesting from this analysis that the overall E_{mw} and AE metrics are about the same for both methods, yet the overall E_m value for the Stork method is about 50 times smaller than that for the Woodward−Rabe method owing to the higher reaction yields and higher atom economies per step.

In the Woodward−Rabe linear sequence, it is evident that the greatest contributors of waste are low reaction yields and low atom economies per step. These contribute together to lower the reaction mass efficiencies. Of the 19 steps, 4 have yields exceeding 75% with an average yield per step of about 65%. The two reactions that contribute the greatest waste in the Woodward−Rabe sequence are the first step which

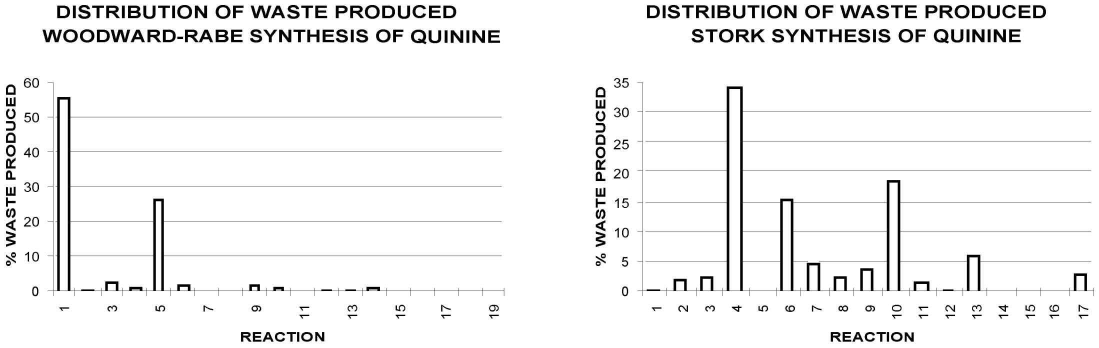


Figure 4. **Distribution of waste profiles for Woodward−Rabe and Stork synthesis of quinine.**

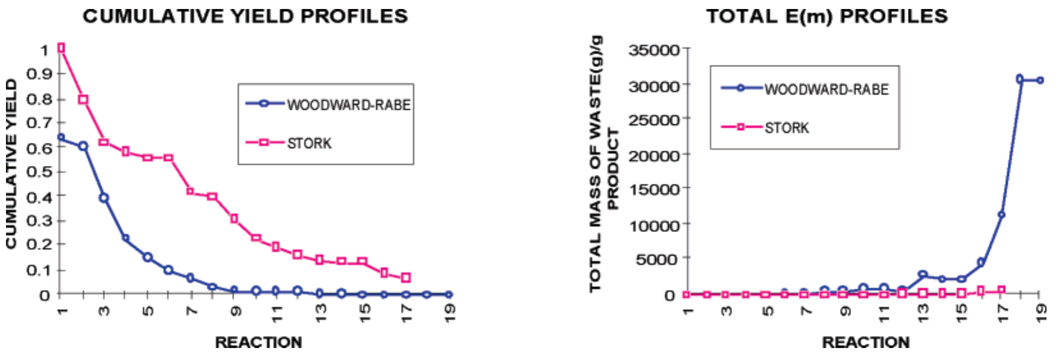


Figure 5. **Cumulative yield and environmental impact factor based on mass profiles for Woodward−Rabe and Stork syntheses of quinine.**

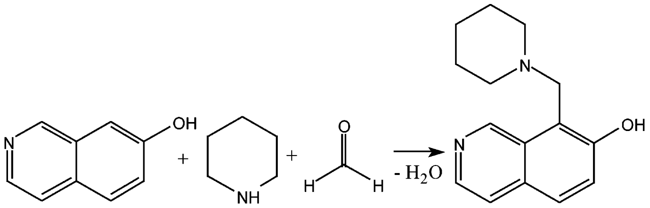
Table 6. **Comparison of the Woodward−Rabe and Stork syntheses of quinine according to various criteria[a]**

criterion	Woodward−Rabe method	Stork method
number of steps	19 (24 actual)	17 (20 actual)
fraction of reactions run under stoichiometric conditions	53%	59%
average yield per step[b]	65%	86%
fraction of reactions run at ambient temperatures	37%	59%
fraction of reactions involving protection deprotection strategies	12.5%	24%
fraction of reactions where byproducts may be recycled back to reagents	37%	82%
fraction of catalytic reactions	47%	24%
fraction of redox reactions	32%	24%
fraction of skeletal building reactions	32%	35%
fraction of MCRs	5%	0%
fraction of substitution reactions	21%	6%
fraction of elimination reactions	32%	35%
fraction of rearrangement reactions	0%	0%

[a] Fractions based on 19 steps for Woodward-Rabe synthesis and 17 steps for Stork synthesis. [b] Geometric mean.

involves conversion of a nitrobenzene to a phenol via reduction, diazotization, and then substitution by the Sandmeyer method and the fifth step which involves loss of a piperidine which represents about 35% loss in molecular weight from the starting material. It is interesting to note that the immediately preceding step to this involved a three-component coupling Mannich-like reaction as shown below in Scheme 4, whereas no multicomponent strategies were employed in the Stork sequence. In the Stork synthesis the greatest contributor of waste is nonstoichiometric reaction conditions particularly for reactions 4, 6, 10, 13, and 17

Scheme 4

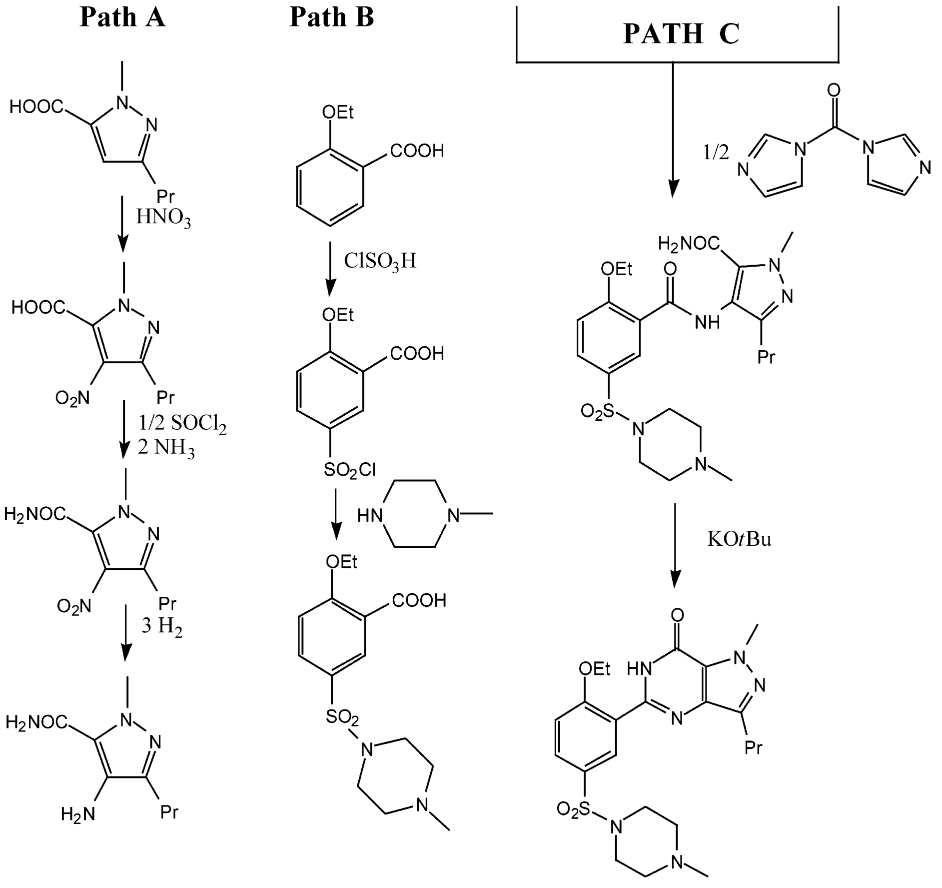


though the reaction yields and atom economies are much improved overall. Though the Stork method produces less mass of waste per mass of target quinine in fewer steps, it involves protection and deprotection steps. This is one of the strategies to be minimized in the 12 Principles of Green Chemistry,[12] and its use or nonuse is directly related to a proposed retrosynthetic plan. If it is not possible to avoid using this technique, then as far as "greenness" is concerned it may still be possible to minimize wastes if protecting groups are carefully chosen so that recycling of liberated groups generated in deprotection steps to appropriate protecting group reagents is shown to be viable and advantageous by an analysis such as that proposed in Table 3. On the point of minimizing protection−deprotection, the Woodward−Rabe synthesis has the edge, since this technique is employed in 3 out of 24 reactions in the plan, whereas the fraction in the Stork plan is 4 out of 17. On the other hand, the Stork method has the greatest potential for recycling products back to reagents. The most important of these is the recycling of triphenylphospine oxide byproduct back to triphenylphosphine reagent by a reduction reaction. This byproduct is formed in three reactions in the Stork plan. Triphenylphos-

---

(12) Anastas, P. T.; Warner, J. C. *Green Chemistry: Theory and Practice*; Oxford University Press: New York, 1998.

Scheme 5



phine oxide is a bulky waste product and is also the major waste product in the Wittig reaction which is commonly criticized as having a low atom economy.[6a,12,13] Such a reduction reaction that is commercially viable has been reported.[14]

8. Application to Convergent Sequences

A simplified algorithm for determining the overall reaction mass efficiency and cumulative mass of waste for a convergent scheme of two linear sequences under the assumption of reclaimed reaction and postreaction solvents can be represented by the following series of steps:

1. Analyze each linear sequence separately as in the previous algorithm given in Section 7.

2. Normalize scales of each sequence by multiplying the smaller scale sequence by a factor given by

$$factor = (x_{\mathrm{large}}/x_{\mathrm{small}})(\epsilon_{\mathrm{overall}}^{\mathrm{largescalesequence}}/\epsilon_{\mathrm{overall}}^{\mathrm{smallscalesequence}})$$

This defines the number of times the smaller scale reaction sequence needs to be repeated to match the scale of the larger scale sequence.

3. Multiply mass of waste of smaller scale sequence by the multiplying factor obtained in step 2.

4. Analyze convergent step and subsequent steps using the scale of final product obtained from the larger-scale sequence as per algorithm for linear sequences.

5. Obtain overall mass of waste by summing masses of waste from larger scale sequence, repeated smaller scale sequence, and convergent sequence. Determine mass of final product collected from convergent sequence.

6. Determine overall yield for entire process from

$$\epsilon_{\mathrm{overall}} = (\prod_{k=1}^{n} \epsilon_k)\epsilon_c(\prod_{k=1}^{m} \epsilon_k)$$

7. Determine

$$E_{\mathrm{m}}^{\mathrm{overall}} = [\sum_{j=1}^{n} \bar{w}_j]/x_{\mathrm{large}}\, p_n \epsilon_{\mathrm{overall}}$$

where $\sum_{j=1}^{n} \bar{w}_j$ is the sum of waste obtained in step 5.

8. Determine $(\mathrm{RME})_{\mathrm{overall}} = 1/(1 + E_{\mathrm{m}}^{\mathrm{overall}})$.

As an illustrative example of this algorithm, the convergent synthesis of sildenafil (Viagra) is analyzed, since detailed experimental procedures for its bulk synthesis are available in the literature.[15] The convergent synthesis is shown in Scheme 5, and Table 7 summarizes the relevant reaction metrics. The present calculation which does not

---

(13) (a) Lancaster, M. In *Handbook of Green Chemistry and Technology*; Clark, J., Macquarrie, D., Eds.; Blackwell Science Ltd.: Oxford, 2002; Chapter 2. (b) Matlack, A. S. *Introduction to Green Chemistry*; Marcel Dekker: New York, 2001.

(14) (a) Hermling, D.; Bassler, P.; Hammes, P.; Hugo, R.; Lechtken, P.; Siegel, H. U.S. Patent 5,527,966, 1996. (b) Lechtken, P.; Sauer, F.; Fankhaenel, M.; Hermling, D. U.S. Patent 5,481,041, 1996.

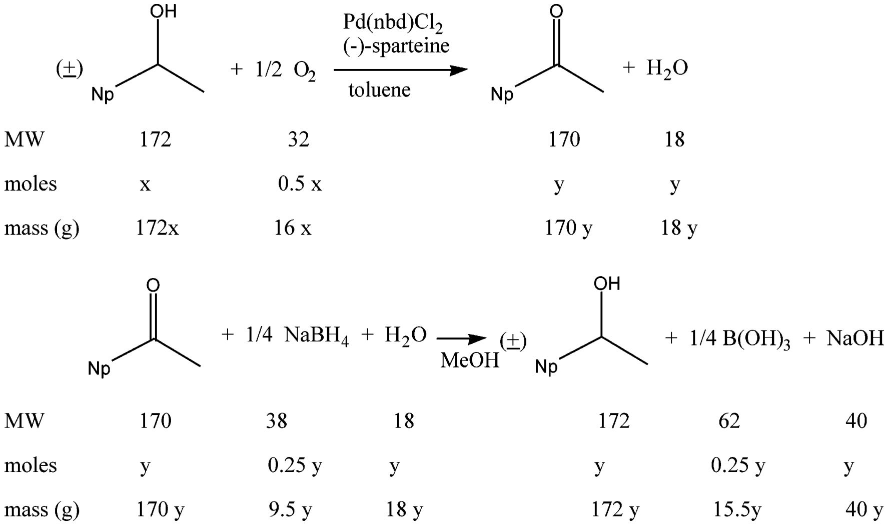
(15) (a) Dunn, P. J.; Galvin, S.; Hettenbach, K. *Green Chem.* **2004**, *6*, 43. (b) Dale, D. J.; Dunn, P. J.; Golightly, C.; Hughes, M. L.; Levett, P. C.; Pearce, A. K.; Searle, P. M.; Ward, G.; Wood, A. S. *Org. Process Res. Dev.* **2000**, *4*, 17.

Table 7. **Summary of kernel reaction metrics analysis for sildenafil (Viagra) synthesis determined by algorithm described in text under the assumptions of recoverable reaction solvents and catalysts, solvents used in workup and purification procedures, and each intermediate product committed as a substrate in the next consecutive step in the sequence**[a]

step no.	reaction type	MW, p_j	ϵ	SF	E_{mw}	AE	RME	E_{m}	mass waste (g)
1A	nitration	213	0.96	1.07	0.085	0.922	0.83	0.21	403.62
2A	amidation	212	0.923	3.53	0.446	0.692	0.18	4.53	8035.70
3A	reduction	182	1	1[a]	0.198	0.835	0.84	0.20	301.14
	overall path A scale of reaction = 9.45 mol		0.886		0.816	0.551	0.15	5.74	8740.46
1B	chlorosulfonation	264.45	0.906	2.30	0.068	0.936	0.37	1.71	61.53
2B	sulfoamination	328	0.86	1.37	0.111	0.900	0.56	0.77	29.52
	overall path B scale of reaction = 0.15 mol		0.779		0.166	0.858	0.30	2.38	91.05
	overall path B repeated 71.54 times								6513.72
1C	amidation[b]	492	0.90[c]	1.24	0.201	0.832	0.60	0.66	2431.22
2C	cyclization[d]	474	0.902	1.04	0.274	0.785	0.68	0.47	1509.21
	overall path C scale of reaction = 8.37 mol		0.812		0.483	0.674	0.45	1.22	3940.43
	overall process		0.72		0.911	0.523	0.144	5.96	19194.61

[a] Assumed. [b] Convergent step. [c] Combined yield for two steps. [d] Forward step.

Scheme 6



include reaction solvents and solvents used in the workup and purification steps affords the calculated overall environmental impact factor based on mass, E_{m}, of 5.96 kg of waste per kg of sildenafil. This is in excellent agreement with and confirms the previously reported value for the overall Sheldon *E*-factor value of 6 kg of waste per kg of product in a process that claimed minimization, elimination, and/or recovery of solvent usage in the manufacture of this pharmaceutical.[15] The overall yield and overall atom economy obtained in this analysis are also in agreement with reported values of 75% and 54%, respectively.

9. Kinetic Resolution of Chiral Substrates

Kinetic resolution of racemic mixtures of chiral substances may be encountered in synthetic plans to stereoisomeric pharmaceutical targets, though such a step is strategically placed downstream in the overall plan so as to minimize the waste of starting materials. The idea is to convert preferentially the unwanted stereoisomer to another compound of different structure so that separation of the resulting new product from the untouched target stereoisomer is facilitated. Recently Stoltz and Ferreira have reported an ingenious variant of kinetic resolution of chiral secondary alcohols which incorporates a recycling reduction step to further amplify the resolution process. In this strategy as shown in Scheme 6, the initial kinetic resolution step involves stereoselective oxidation of (+) 1-naphthylethanol to naphthylmethyl ketone leaving the target (−) alcohol untouched. After separation of the (−) alcohol the collected ketone is subjected to a reduction back to the racemic alcohol in near quantitative yield. The process is repeated a number of times until a desired amount of target (−) alcohol is reached.

Generally, if λ_1 is the fraction of starting racemic alcohol converted to ketone, $1 - \lambda_1$ is the fraction of unreacted racemic alcohol, and λ_2 is the fraction of unreacted racemic

Table 8. **Summary of results for preferential resolution of (±)1-naphthylethanol to (−)1-naphthyethanol by successive stereoselective oxidation of (+)1-naphthylethanol to naphthylmethyl ketone and reduction of the ketone back to (±)1-naphthylethanol**

cycle	mass of (−)isomer (g)[a]	mass of (+)isomer (g)[a]	RME[b]	% gain in RME[c]	E_m (g/g)	ω(max)[d]
1	2.20	2.80	0.373		1.681	
2	3.41	1.59	0.533	42.9	0.876	210.22x
3	4.08	0.92	0.630	68.9	0.639	127.78x
4	4.44	0.56	0.651	74.5	0.536	88.46x
5	4.64	0.36	0.672	80.2	0.488	66.09x
6	4.75	0.25	0.683	83.1	0.463	52.00x
7	4.81	0.19	0.690	85.0	0.450	42.48x
8	4.85	0.15	0.693	85.8	0.443	35.72x
9	4.87	0.13	0.695	86.3	0.439	30.71x
10	4.88	0.12	0.696	86.6	0.437	26.88x
limit	4.89	0.11	0.697	86.9	0.434	0

[a] Initial mass of racemic alcohol is 5 g. [b] Cumulative reaction mass efficiency assuming that all reaction and postreaction solvents are recovered. [c] With respect to RME = 0.373 in cycle 1. [d] Maximum mass of all solvents and other ancillary materials in reaction and postreaction procedures for both oxidation and reduction steps that is allowed to be discarded for a net positive gain in RME with respect to initial value of RME in cycle 1; x is the number of moles of starting racemic alcohol that defines the scale of the entire resolution operation.

alcohol that is the (−) stereoisomer, then according to Scheme 6 the masses of (−) and (+) alcohols after the oxidation step are $172x\lambda_2(1-\lambda_1)$ and $172x(1-\lambda_1)(1-\lambda_2)$, respectively. Since $\lambda_1 x$ is the number of moles of starting racemic alcohol that is converted to ketone, then the number of moles of collected ketone after oxidation, y, is also equal to $\lambda_1 x$. Subsequent quantitative reduction of all of this ketone yields $172x\lambda_1$ grams of regenerated racemic alcohol. Under stoichiometric conditions for both reactions, the RME with respect to the desired (−) alcohol after the first oxidation−reduction cycle is given by

$$\mathrm{RME}_1^{(-)} = \frac{172x(1-\lambda_1)\lambda_2}{x(188+27.5\lambda_1)+\omega} \tag{43}$$

where the denominator accounts for the masses of initial racemic alcohol starting material and of the subsequent sodium borohydride and water used in the recycling reduction step and ω represents the mass of all reaction solvents (toluene and methanol) and postreaction solvents and materials used in both the oxidation and reduction steps.

After m oxidation−reduction cycles, the cumulative masses of (+) and (−) alcohol are given by $172x(1-\lambda_1{}^m)(1-\lambda_2)$ and $172x(1-\lambda_1{}^m)\lambda_2$, respectively; and the cumulative RME with respect to the desired (−) alcohol is given by

$$\mathrm{RME}_m^{(-)} = \frac{172x(1-\lambda_1)(1-\lambda_1^m)\lambda_2}{x[(16+27.5\lambda_1)(1-\lambda_1^m)+172(1-\lambda_1)]+m(1-\lambda_1)\omega} \tag{44}$$

(see Supporting Information for derivation). The underlying assumption is that the values of λ_1 and λ_2 remain constant from cycle to cycle and are independent of the scale of the reactions.

Table 8 summarizes results of this analysis applied to a starting mass of racemic 1-naphthylethanol of 5 g ($x = 5/172 = 0.0291$ mol) with $\lambda_1 = 0.55$ and $\lambda_2 = 0.98$ taken from the reports of Stoltz and Ferreira.[16] The entries under the RME and E_m columns are calculated according to eq 44 with ω set to zero (meaning all reaction and postreaction solvents and materials other than reagents are recovered) and the reciprocal relationship in eqs 7, 17, or 31. These results represent best-case scenarios. It is evident that with each succeeding cycle the resolution improves. From a practical point of view, the greatest gains in RME are achieved in the first three cycles after an initial value of 37.3%. Thereafter, the gain in RME is incremental and plateaus to a limiting value of 69.7%. From the general relation in eq 44, which incorporates masses of reaction and postreaction solvents, it is possible to determine a threshold or maximal value of ω so that the fractional gain in RME for cycle m relative to the first cycle is positive. Such a fractional gain is given by

$$\Delta_{\mathrm{RME}} = \frac{\mathrm{RME}_m^{(-)}}{\mathrm{RME}_1^{(-)}} - 1 = \frac{172x(\Omega-1)-\omega(m-\Omega)}{x(31.125\Omega+172)+m\omega} \tag{45}$$

where

$$\Omega = \sum_{j=0}^{m-1}\lambda_1^j = \sum_{j=0}^{m-1}(0.55)^j$$

and $m \geq 2$. For $\Delta_{\mathrm{RME}} > 0$, the mass of reaction and postreaction solvents must not exceed $172x(\Omega-1)/(m-\Omega)$. This maximum value of ω sets an upper limit for the masses of solvents and other ancillary materials that can be used in both operations under the provision that these will *not* be recovered, as noted in section 3 and Table 1, and still result in a positive gain in cumulative RME for the desired product. In essence this is really saying how much ancillary reaction material can be afforded to be considered as net waste in the entire process. A further simplification in this analysis is that solubilities of materials are invariant with the magnitude of the scale x which may or may not hold true. For the present illustration the last column in Table 8 lists the threshold maximal values of ω for each oxidation−reduction cycle after the first cycle. For example, if $x = 0.0291$ mol, corresponding to 5 g of racemic alcohol starting material, then for two cycles $\omega_{\max} = 210.22(0.0291) = 6.1$. Hence, no more than 6.1 g of combined reaction and postreaction solvents and ancillary materials, which are destined to be discarded, must be used in both operations in order to obtain a net increase in cumulative RME for the process after two cycles. It is evident that when several cycles are attempted in the resolution it becomes more imperative to recover solvents and/or minimize their usage as the maximal threshold value of ω decreases rapidly. Within the constraints imposed, the present analysis clearly demonstrates that recovery, minimization, or elimination of ancillary materials in reaction processes appears to be crucial in meeting the objective of high reaction mass efficiencies.

---

(16) (a) Ferreira, E. M.; Stoltz, B. M. *J. Am. Chem. Soc.* **2001**, *123*, 7725. (b) Stoltz, B. M. *Chem. Lett.* **2004**, *33*, 362.

Conclusions

This work provides a unified protocol for the use of various reaction metrics for the analysis of "greenness" of chemical reactions. A new reaction metric, stoichiometric factor, is introduced which takes into account reactions run under nonstoichiometric conditions. A general expression for reaction mass efficiency is derived, and it is shown to depend on four competing factors: reaction yield, atom economy, stoichiometric factor, and a factor accounting for reaction and postreaction solvent and/or catalyst recovery. Simple algorithms that can be implemented in spreadsheet formats have been presented, and their use has been demonstrated for individual reactions that produce single products or isomeric products, linear sequences of reactions, convergent sequences of reactions, and kinetic resolution of chiral substrates. Conditions for advantageous recycling and recovery of solvents, catalysts, and/or byproducts have also been examined within the constraint of raw materials consumption. Raw materials costs have also been included in the analyses. Solvent recovery or elimination in the reaction and postreaction operations appears to be the most critical factor in maintaining high overall reaction mass efficiencies.

Acknowledgment

Parts of this work were presented at the Canada−U.S. Joint Workshop on Innovative Chemistry in Clean Media, May 20−21, 2004 at McGill University, Montreal, Canada and at the 39th IUPAC Congress and 86th Canadian Society of Chemistry Conference & Exhibition, August 10−15, 2003 in Ottawa, Canada. York University is thanked for a Professional Development Fund grant in support of this work. The helpful comments of an anonymous referee in chemical process R&D are greatly appreciated.

Supporting Information Available

Supporting Information contains reaction schemes for Woodward−Rabe and Stork synthetic routes to quinine showing balanced chemical equations for each step that form the basis of results given in Table 4 and derivations of eqs 7, 17, 31, and 44. This material is available free of charge via the Internet at http://pubs.acs.org.

NOMENCLATURE

a, molecular weight of reaction substrate

AE, Trost atom economy

b, molecular weight of reagent

c, mass of catalyst

ϵ, reaction yield

$\langle\epsilon\rangle$, composite reaction yield corresponding to isolated target product after reaction, workup, and purification steps

ϵ^*, reaction yield for recycling reaction

E_{mw}, environmental impact factor based on molecular weight assuming 100% reaction yields

E_m, environmental impact factor based on mass that at least takes into account reaction yield

E, Sheldon environmental impact factor taking into account all possible forms of waste produced

ϕ, excess moles of reagent used

f_1, ratio of masses of catalysts and all solvents used in product forming reaction to mass of target product

f_2, ratio of masses of catalysts and all solvents used in product forming and recycling reactions to mass of target product

f_j, fraction of jth isomeric product in product mixture

M, mass of target product P in a linear or convergent sequence

n, number of steps in a linear sequence

p, molecular weight of target product

q, molecular weight of reaction byproduct

RMC, actual raw materials cost

RMC_{100}, raw materials cost assuming 100% reaction yield

RME, reaction mass efficiency

$\$_A$, cost per unit mass of reagent A

s, mass of reaction solvent

SF, stoichiometric factor taking into account reagents used in excess: $SF = 1$ for stoichiometric reactions, $SF > 1$ for nonstoichiometric reactions

x, moles of reagents used (stoichiometric conditions)

y, moles of target product after reaction

y'', moles of target product isolated after reaction, workup, and purification

w, mass of unreacted starting materials

$\bar{w}$, mass of total waste produced in a reaction

z, moles of reagents used (nonstoichiometric conditions)

ω, mass of all materials used in postreaction operations: workup, purification, waste stream treatment

Received for review November 2, 2004.

OP049803N