

Unification of Reaction Metrics for Green Chemistry II: Evaluation of Named Organic Reactions and Application to Reaction Discovery

John Andraos*

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

Abstract:

The concept of minimum atom economy (AE_{min}) and maximum environmental impact factor E_{max} is introduced and applied to a database of more than 400 named organic reactions by employing generalized Markush structures as a means of gauging worst-case scenarios for waste production for chemical reactions. The scope of the method can be extended to evaluate “green” performances for any chemical reaction once all byproducts are identified. From the inverse relationship connecting AE and E_{mw} (environmental impact factor based on molecular weight) and an analogous one connecting RME (reaction mass efficiency) and E (Sheldon environmental impact factor), a minimum value of AE or RME equal to the golden ratio, 0.618 (or 61.8%), is suggested as a threshold metric for gauging the true “greenness” of reactions. The rationale for this is that this condition is met when the magnitude of AE exceeds that of E_{mw} and similarly when RME exceeds E . Probabilities for achieving this condition are also determined for several organic reactions, and general expressions for evaluating such probabilities as functions of a general threshold α between 0 and 1 are also derived. Reactions in the database are classified according to general reaction types, and each class is ranked according to the “golden” threshold metric. Patterns with respect to atom economical efficiency revealed by this analysis are discussed, including the introduction of expressions for determining the viability of recycling waste byproducts back to reagents. It is shown that the database can be used in a diversity-oriented approach to discover new multicomponent reactions (MCRs) by combinatorial analysis of Markush fragments. In this context the top seven most frequently appearing Markush structures in the database yield 33 viable three-component MCRs of which 12 have literature precedence. Synthetic strategies for reaction design are put forward using the optimum criteria suggested by analysis of the database.

1. Introduction

A unified approach for determining various reaction metrics that can be used to gauge the “greenness” of organic reactions has recently been described.¹ Key algebraic expressions, given in eqs 1–3, were found that linked reaction yield (ϵ), Trost’s atom economy (AE), environmental impact factor based on molecular weight (E_{mw}), Sheldon’s environmental impact factor based on mass ($E_m = E$), reaction mass efficiency (RME), stoichiometric factor (SF) accounting for

reactions run with excess reagents, and a solvent and catalyst environmental impact parameter (f).

$$AE = \frac{1}{1 + E_{mw}} \quad (1)$$

$$RME = \frac{1}{1 + E_m} \quad (2)$$

$$RME = \begin{cases} \frac{\epsilon(AE)}{SF}, & \text{recovery of all solvents and catalysts} \\ \frac{\epsilon(AE)}{SF} \left[\frac{1}{1 + f \frac{\epsilon(AE)}{SF}} \right], & \text{no recovery of solvents and catalysts} \end{cases} \quad (3)$$

where $0 \leq AE \leq 1$, $0 \leq RME \leq 1$, $0 \leq \epsilon \leq 1$, $SF = 1$ for stoichiometric reactions, $SF > 1$ for nonstoichiometric reactions, $E_{mw} > 0$, $E_m > 0$, and $f = 0$ if all solvents, catalysts, and other materials used in reaction and postreaction processes are recovered, recycled, and/or eliminated, otherwise, $f > 0$. For a balanced chemical equation the AE and E_{mw} metrics are determined using molecular weights of reagents and products, whereas SF and f are determined using actual masses of materials used according to

$$SF = 1 + \frac{(AE) \sum \text{mass of excess reagents}}{\text{mass of target product at 100% yield}} \quad (4)$$

$$f = \frac{\sum \text{mass of reaction and postreaction solvents and materials} + \text{mass of catalysts used}}{\text{mass of collected target product}} \quad (5)$$

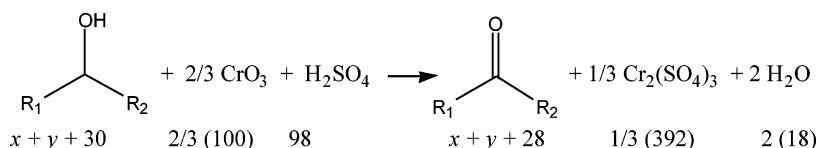
Various applications were presented which showed how these parameters could be used in determining overall waste production for individual reactions that produce single products or isomeric products. Also considered were simple linear sequences of reactions, simple convergent sequences of reactions, and product forming reactions coupled with a recycling reaction that converts a byproduct back to a starting material. Conditions for advantageous recycling and recovery of solvents, catalysts, and/or byproducts were also examined within the constraint of raw materials consumption. Conclusions drawn for the optimization of syntheses of useful targets are in agreement with a recent analysis based on cumulative atom economies presented by Eissen.²

The present work now extends these ideas to generalized chemical reactions so that worst-case scenarios for atom

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

(1) Andraos, J. *Org. Process Res. Dev.* 2005, 9, 149.

Scheme 1



economy and waste production may be determined. For example for a given chemical reaction, it is possible to ask what is the lowest possible atom economy or the highest possible environmental impact factor that might be expected. We begin by generalizing reactions by writing out balanced chemical equations accounting for all byproducts using Markush structures for substrates, reagents, and products so that variable chemical groups or atoms are distinguished from those that need to be specified. In any kind of chemical transformation some fragment of the substrate chemical structure is operated upon, whereas other parts remain untouched. The relevant atoms involved in the transformation are thus specified while the spectator atoms are designated with the customary R group notation. From this Markush-type chemical equation, general expressions for AE and E_{mw} as functions of R group size may be obtained under the assumption of 100% reaction yield, stoichiometric operating conditions, and complete recovery or elimination of solvents. It is then possible to estimate the minimum atom economy, $(\text{AE})_{\text{min}}$, or the maximum unit waste expected per unit target product, E_{max} , according to the dependence of these metrics on R group size. These estimates of course represent “best-case” scenarios since we already know that once these constraints are lifted the maximum E_{mw} , E_{max} , and the minimum AE, $(\text{AE})_{\text{min}}$, estimates will increase and decrease, respectively. In fact, the maximum E_{mw} becomes the maximum E , and the minimum AE becomes the minimum RME. It is argued that in order to have the best chance to achieve “greenness” the kernel reaction metrics E_{mw} and AE derived from the intrinsic nature of a given chemical reaction need to be optimized from the outset. This in turn is directly linked to the synthetic design or plan of a given target product. Also evident from the inverse relationships in eqs 1 and 2 is that maximization of $E_{\text{mw}}(E)$ automatically implies minimization of AE (RME), and vice versa. Hence, estimation of an upper bound for $E_{\text{mw}}(E)$ coincides with estimation of a lower bound for AE (RME).

To better quantify what “greenness” is we introduce the concept of a threshold criterion. If a reaction is found to have an environmental impact factor below a certain threshold that defines “greenness”, then it is possible to rank the “greenness” of the reaction using this criterion against others that are of the same or different chemical class.

Recognizing that $0 \leq \text{AE} \leq 1$ we may set a threshold value of α between 0 and 1 such that a reaction is considered “green” if $\text{AE} \geq \alpha$. From eq 1 we would then have a corresponding environmental impact factor based on molecular weight threshold of $E_{\text{mw}} \leq (1/\alpha) - 1$. Similarly, following the same argument, we obtain from eq 2 the thresholds $\text{RME} \geq \alpha$ and $E \leq (1/\alpha) - 1$, respectively. It is shown that an appropriate threshold value for “greenness” is $\alpha = 0.618$ for both the estimates of upper bounds for E_{max}

and lower bounds for $(\text{AE})_{\text{min}}$. This means that a chemical reaction satisfies the “green” criterion if $(\text{AE})_{\text{min}}$ is at least 61.8% or if E_{max} is at most 0.618 mass unit of waste per mass unit of target product on a 1 mol scale. It is argued that satisfying this criterion for the kernel metrics AE and E_{mw} is the key to increasing the probability that this same threshold is met when considering the more realistic metrics RME and E which include reaction yield, the use of excess reagents, and reaction and postreaction solvents and catalysts. If the threshold criterion is not met at the kernel metric level, then there is no hope that it will be met when these other parameters are taken into account. The probabilities of satisfying this criterion as functions of the threshold α are precisely determined for four case scenarios: (a) stoichiometric reactions with complete solvent recovery and/or elimination in the reaction and postreaction steps, (b) nonstoichiometric reactions with complete solvent recovery and/or elimination in the reaction and postreaction steps, (c) stoichiometric reactions with no recovery of solvents and/or catalysts, and (d) nonstoichiometric reactions with no recovery of solvents and/or catalysts.

Using compilations of well-known named organic reactions a Markush database of organic reactions was generated, and $(\text{AE})_{\text{min}}$ and E_{max} were determined for each reaction. Patterns and rankings among reaction classes are discussed including an analysis of the viability of recycling waste byproducts back to reagents. Simple frequency and combinatorial analyses of Markush structures allow the database to be used as a toolbox to discover and predict new multicomponent reactions in a diversity oriented approach. In this regard the top seven most frequently occurring Markush structures in the database are used to construct 33 viable one-pot, three-component coupling reactions. A literature search indicates that 12 of these have recently been verified experimentally.

2. Concept of Minimum and Maximum Reaction Metrics

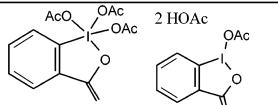
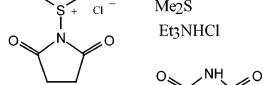
The concept of minimum AE and maximum E_{mw} is illustrated with the generalized form of the Jones oxidation of alcohols shown in Scheme 1 run under stoichiometric conditions with complete recovery of reaction and postreaction solvents. Markush structures are used for the starting alcohol and product carbonyl compound showing both variable R groups and necessarily invariant atoms.

If x and y represent the molecular weights of fragments R_1 and R_2 , then the general expressions for E_{mw} and AE are given by eqs 6 and 7.

$$E_{\text{mw}} = \frac{(1/3)392 + (2)18}{x + y + 28} = \frac{500}{3(x + y + 28)} \quad (6)$$

$$\text{AE} = \frac{x + y + 28}{x + y + 28 + (1/3)392 + (2)18} = \frac{3(x + y + 28)}{3(x + y) + 584} \quad (7)$$

Table 1. Summary of minimum AE and maximum E_{mw} parameters for various oxidations of alcohols under stoichiometric conditions

Reaction	Oxidizing Agent	Byproducts	AE ^a (min)	E ^a (max)	AE ^b (min)	E ^b (max)
Dess-Martin			0.062	15.21	0.12	7.34
Swern	$\text{Me}_2\text{S}=\text{O}$ Et_3N	Me_2S Et_3NHCl	0.083	11	0.16	5.33
Corey-Kim	$\text{Cl}(\text{CO})(\text{CO})\text{Cl}$ 	$\text{Cl}(\text{CO})\text{COOH}$ Me_2S Et_3NHCl	0.086	10.66	0.16	5.15
Pfitzner-Moffatt	$\text{Me}_2\text{S}=\text{O}$ DCC^c	Me_2S	0.09	10.14	0.17	4.90
Dichromate	1/2 $\text{K}_2\text{Cr}_2\text{O}_7$ 1/2 H_2SO_4	H_2CrO_3 1/2 H_2O 1/2 K_2SO_4	0.13	6.60	0.23	3.41
Jones	2/3 Cr_2O_7 H_2SO_4	$\text{Cr}_2(\text{SO}_4)_3$ 6 H_2O	0.15	5.56	0.26	2.87
Permanganate	2/3 KMnO_4 1/3 H_2SO_4	2/3 MnO_2 4/3 H_2O 1/3 K_2SO_4	0.18	4.67	0.29	2.41
Hydrogen peroxide oxidation	H_2O_2	2 H_2O	0.45	1.20	0.62	0.62
Uemura	1/2 O_2	H_2O	0.63	0.60	0.76	0.31
Theoretical limit	????	H_2	0.93	0.075	0.97	0.034

^a Calculations based on kernel reaction of oxidation of methanol to formaldehyde. ^b Calculations based on kernel reaction of oxidation of 2-propanol to acetone. ^c Dicyclohexylideneimide.

Of the two metrics it is easier to determine the limiting values of E_{mw} . Once these have been established, the limiting values of AE are at once determined using the connecting relationship given by eq 1. It is obvious that the value of E_{mw} decreases as the molecular weights of R_1 and R_2 increase. Thus, as x and y both become large, the value of E_{mw} approaches zero, and this defines the minimum value, E_{min} . If R_1 and R_2 are both as small as possible, that is, if they represent hydrogen atoms, then the largest possible value of E_{mw} for this reaction is $E_{max} = [500/3(1 + 1 + 28)] = 5.56$. Hence, the range of E_{mw} for the Jones oxidation of primary alcohols to aldehydes is $0 < E_{mw} < 5.56$. Applying this range to eq 1 yields the range of values of AE. Therefore, $1 > AE > 0.153$ where

$$(AE)_{min} = \frac{1}{1 + E_{max}} \quad (8)$$

and

$$(AE)_{max} = \frac{1}{1 + E_{min}} \quad (9)$$

apply. We may conclude that the minimum atom economy for the Jones oxidation of primary alcohols to aldehydes is about 15.3% and that the maximum waste generated is 5.56 mass units per mass unit of aldehyde on a 1 mol scale, where the kernel reaction is taken to be the oxidation of methanol to formaldehyde. Similar analysis using secondary alcohols as substrates yields a minimum AE of 25.8% and a maximum E_{mw} of 2.87 mass units of waste per mass unit of ketone

Scheme 2

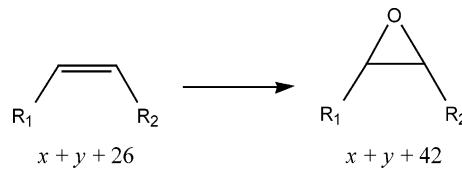
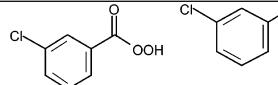


Table 2. Summary of minimum AE and maximum E_{mw} parameters for various epoxidations of olefins under stoichiometric conditions

Reaction	Oxidizing Agent	Byproducts	AE(min) ^a	E(max) ^a
Prilezhaev		$\text{Cl}-\text{C}_6\text{H}_4-\text{COOH}$	0.22	3.56
Sharpless	$tert\text{-BuOOH}$	$tert\text{-BuOH}$	0.37	1.68
Jacobsen	NaOCl	NaCl	0.43	1.33
Hydrogen peroxide	H_2O_2	H_2O	0.71	0.41
Theoretical limit	$1/2 \text{O}_2$	None	1	0

^a Calculations based on kernel reaction of epoxidation of ethylene to ethylene oxide.

product on a 1 mol scale, where the kernel reaction is taken to be the oxidation of 2-propanol to acetone. Table 1 summarizes various methods of oxidizing primary and secondary alcohols to aldehydes and ketones, respectively. One may be able to rank at once the “green” performance of the Jones oxidation against other oxidation methods operating on the same substrates by examining the magnitudes of the $(AE)_{min}$ and E_{max} parameters in each case. For this kind of transformation it is obvious that the smallest possible waste byproduct is hydrogen gas. Indeed, this observation defines the theoretical upper and lower limits of $(AE)_{min}$ and E_{max} , respectively, for this reaction as shown in Table 1.

Another example of this kind of analysis is demonstrated for the epoxidation of olefins as shown in Scheme 2 and Table 2. For this kind of reaction the theoretical limit of generating no waste has been achieved as evidenced by the well-known industrial process to produce ethylene oxide from ethylene with a silver catalyst.³

3. Reactions Yielding More than One Useful Product

In the previous examples, e.g., eq 6, the reactions produced waste byproducts that did not contain any of the variable R groups in their structures. This led to expressions of E_{mw} having the general form

$$E_{mw} = \frac{c_1}{c_2x + c_3} \quad (10)$$

where c_1 , c_2 , and c_3 are constants and x is the molecular weight of the R group fragment. Figure 1a illustrates a graphical representation of this function that shows that E_{mw}

(3) (a) Lefort, T. E. (Société Francaise de Catalyse Generalisée). FR 729,952, 1931. (b) Lefort, T. E. (Société Francaise de Catalyse Generalisée). FR 739,- 562, 1931.

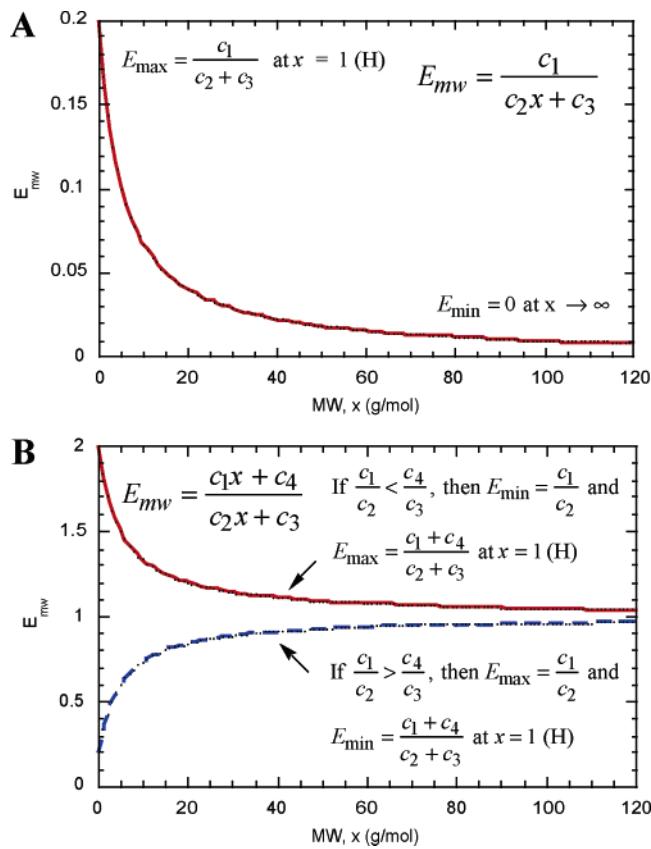


Figure 1. Behaviour of general form of environmental impact factor E_{mw} , for (A) reactions having byproducts not containing variable substituents ($c_1 = 20$, $c_2 = 20$, $c_3 = 100$) and (B) reactions having byproducts containing variable substituents ($c_1 = 20$, $c_2 = 20$, $c_3 = 100$, and $c_4 = 200$ for decay curve; $c_1 = 20$, $c_2 = 20$, $c_3 = 100$, and $c_4 = 20$ for growth curve).

is a monotonically decreasing function of x and reaches a minimum value of zero for very large x values (large substituents) and a maximum value of

$$E_{\max} = \frac{c_1}{c_2 + c_3}$$

for $x = 1$ corresponding to $R = H$. When reactions yield more than one possibly useful product so that both contain variable substituents as shown by the examples in Scheme 3, the expression of E_{mw} takes on the general form given by eq 11.

$$E_{mw} = \frac{c_1x + c_4}{c_2x + c_3} \quad (11)$$

In these situations the optimization of E_{mw} depends on which product is targeted. Figure 1b illustrates the graphical behaviour of this function where it is evident that E_{mw} can be either a monotonically decreasing or increasing function of x , depending on the relative magnitudes of the ratios (c_1/c_2) and (c_4/c_3). Table 3 summarizes the optimization scenarios for E_{mw} corresponding to the reactions shown in Scheme 3. It is evident that for each reaction minimization of E_{mw} with respect to one product is achieved in one direction of R group size while minimization of E_{mw} with respect to the other product is achieved in the opposite direction of R group size.

Table 3. Summary of optimization strategies for reactions producing more than one target product

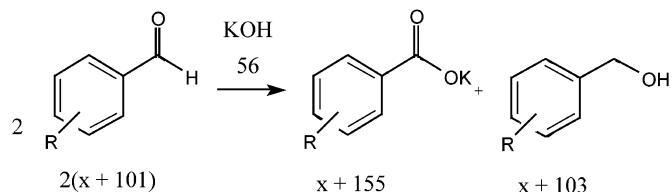
Reaction	Target Product	E_{mw}	Optimization (minimization of E_{mw})
Cannizzaro	carboxylate	$\frac{x+103}{x+155}$	minimized by small R groups
	alcohol	$\frac{x+155}{x+103}$	minimized by large R groups
Hydroperoxide rearrangement	ketone	$\frac{x+17}{2x+28}$	minimized by large R groups
	alcohol	$\frac{2x+28}{x+17}$	minimized by small R groups
Lemieux-Johnson oxidative cleavage of olefins	$R_1\text{CHO}$	$\frac{3y+927}{3x+87}$	minimized by large R_1 and small R_2
	$R_2\text{CHO}$	$\frac{3x+927}{3y+87}$	minimized by small R_1 and large R_2
Harries ozonolysis with oxidative workup	$R_1\text{COOH}$	$\frac{y+63}{x+45}$	minimized by large R_1 and small R_2
	$R_2\text{COOH}$	$\frac{x+45}{y+63}$	minimized by small R_1 and large R_2
Harries ozonolysis with reductive workup	$R_1\text{CHO}$	$\frac{y+107}{x+29}$	minimized by large R_1 and small R_2
	$R_2\text{CHO}$	$\frac{x+107}{y+29}$	minimized by small R_1 and large R_2
Lossen rearrangement	$R_1\text{NH}_2$	$\frac{y+89}{x+16}$	minimized by large R_1 and small R_2
	$R_2\text{COOH}$	$\frac{x+60}{y+45}$	minimized by small R_1 and large R_2
Ritter reaction	amine	$\frac{z+75}{x+y+43}$	minimized by small R' and large R_1 and R_2 groups
	carboxylic acid	$\frac{x+y+43}{z+75}$	minimized by large R' and small R_1 and R_2 groups
Barbier-Wieland reaction	carboxylic acid	$\frac{2y+302.4}{x+45}$	minimized by small R' and large R groups
	ketone	$\frac{x+319.4}{2y+28}$	minimized by large R' and small R groups
Criegee glycol cleavage reaction	ketone (R groups)	$\frac{2y+473}{2x+28}$	minimized by small R' and large R groups
	ketone (R' groups)	$\frac{2x+473}{2y+28}$	minimized by large R' and small R groups

This unique property may be used to advantage in such industrially important reactions as the Hock hydroperoxide rearrangement of substituted cumylhydroperoxides to substituted phenols and ketones as shown in Scheme 4.⁴ For the case of different substituents R_1 and R_2 , if phenols are the target products, then minimization of E_{mw} (maximization of AE) is achieved with small R_2 and large R_1 groups; and if ketones are the target products, then minimization of E_{mw}

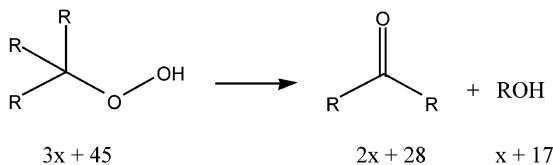
(4) (a) Hock, H.; Lang, S. *Chem. Ber.* **1944**, *77B*, 257. (b) Vandenberg, E. J. (Distillers Co. Ltd.). GB 676,771, 1952. (c) Hulse, G. E.; Vandenberg, E. J. (Hercules Powder Co.). U.S. Patent 2,557,968, 1951.

Scheme 3

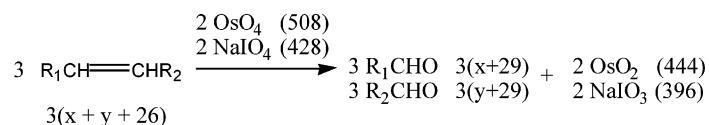
Cannizzaro reaction



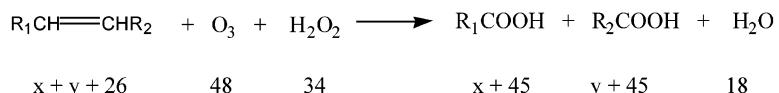
Hydroperoxide rearrangement



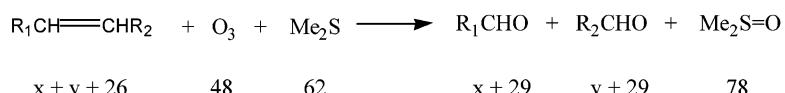
Lemieux-Johnson oxidative cleavage of olefins



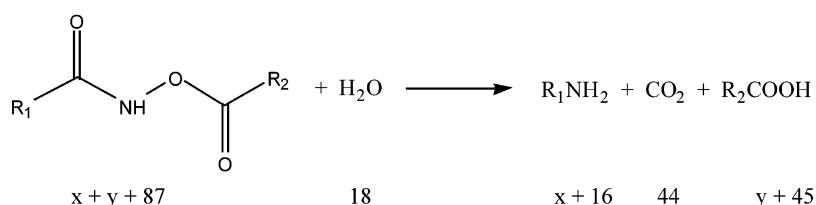
Harries ozonolysis with oxidative workup



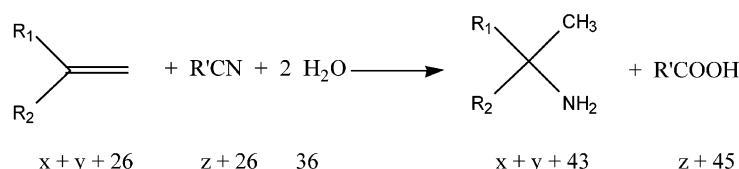
Harries ozonolysis with reductive workup



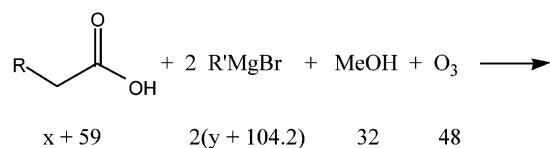
Lossen rearrangement



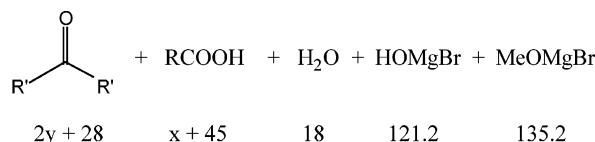
Ritter reaction + hydrolysis



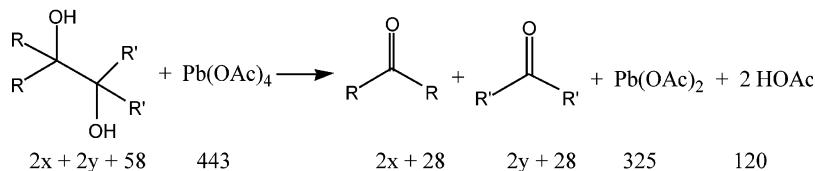
Barbier-Wieland reaction



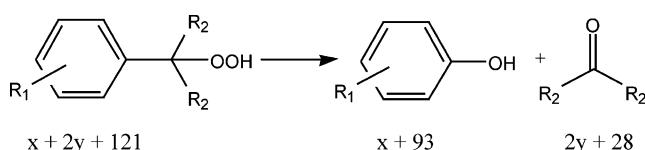
Scheme 3 (Continued)



Criegee glycol cleavage reaction



Scheme 4



$$\text{Phenol target: } E_{\text{mw}} = \frac{2y + 28}{x + 93} \quad \text{Ketone target: } E_{\text{mw}} = \frac{x + 93}{2y + 28}$$

(maximization of AE) is achieved with large R_2 and small R_1 groups. For the case of identical substituents R_1 and R_2 , if phenols are the target products, then $(30/94) < E_{\text{mw}} < 2$ and $0.33 < AE < 0.76$ so that small substituents maximize AE and minimize E_{mw} . If ketones are the target products, then $(1/2) < E_{\text{mw}} < (94/30)$ and $0.24 < AE < 0.67$ so that large substituents maximize AE and minimize E_{mw} . Note that for the case of identical substituents E_{mw} has both lower and upper bounds unlike the former case of different R groups where E_{mw} has no upper bound limit. Consequently, the corresponding range of AE is narrower than the full domain between zero and unity. If the reaction yield is taken into account, then the AE ranges are further attenuated according to $0.33\epsilon < \text{RME} < 0.76\epsilon$ and $0.24\epsilon < \text{RME} < 0.67\epsilon$, respectively where $\text{RME} = \epsilon(\text{AE})$ under the condition of stoichiometric reaction conditions and complete recovery and/or elimination of reaction and postreaction solvents and recovery of any catalysts used.

4. Concept of “Golden” Atom Economy Criterion, Probabilities, and Application to Reaction Database

A graphical analysis of the connecting relationship between AE and E_{mw} given by eq 1 indicates how one may define a threshold for “greenness” for chemical reactions. Figure 2a shows a plot of AE versus E_{mw} in which the graph is divided into two distinct regions, one in which the magnitude of AE exceeds E_{mw} and the other in which the opposite is true. By setting AE equal to E_{mw} in eq 1 one finds that the magnitudes of the atom economy and environmental impact factor based on molecular weight match at a value given by $(\sqrt{5}-1)/2 = 0.61803\dots$. This number, commonly designated by the symbol ϕ , is recognizable as the “golden” ratio and has been known since ancient times.⁵ In the present context this number can be used as a benchmark to identify truly atom economical chemical reactions.

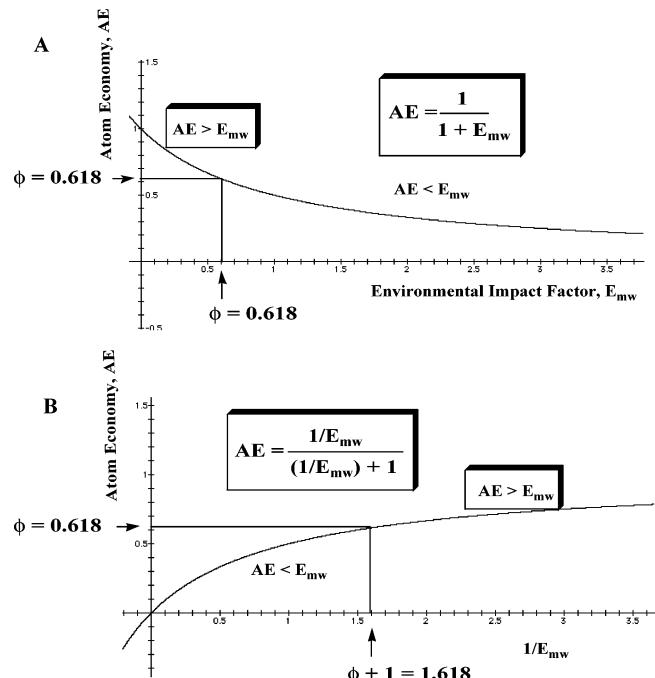


Figure 2. Plots showing the relationship between (A) AE and E_{mw} , and between (B) AE and $1/E_{\text{mw}}$. In each plot the two regions $\text{AE} > E_{\text{mw}}$ and $\text{AE} < E_{\text{mw}}$ are marked.

This means that a true “green” reaction is one which produces at least 0.618 mass units of target product per mass unit of all reactants used ($\text{AE} > 0.618$), or conversely, one that produces at most 0.618 mass units of waste byproducts per mass unit of target product ($E_{\text{mw}} < 0.618$). It should be emphasized that this threshold arises directly from the inverse relationship between AE and E_{mw} that in turn arises directly from a simple analysis of a balanced chemical equation¹ and therefore is not an arbitrary designation. Figure 2b shows a plot of an alternative form of eq 1 that allows, as will be

(5) (a) Walser, H. *The Golden Section*; Mathematical Association of America: Washington, DC, 2001. (b) Livio, M. *The Golden Ratio: The Story of Phi, the World's Most Astonishing Number*; Broadway Books: 2002. It is commonly stated that it is prominent in Nature and has been used as a key geometrical and design element in architecture and Art because of the resultant “beautiful” forms that are created. Cited examples include the Great Pyramid in Egypt and the Parthenon in Athens, Greece, and well-known paintings such as Leonardo da Vinci’s *Mona Lisa* (1504) and Rembrandt van Rijn’s self-portrait (1640). However, many of these claims have proven to be false (see Markowsky, G. *College Math. J.* **1992**, 23, 2 and Falbo, C. *College Math. J.* **2005**, 36, 123).

seen *vide infra*, a more convenient way of plotting AE and E_{mw} data.

In terms of reaction mass efficiency, the “golden” threshold for a single step reaction is given by

$$RME = \epsilon(AE) \geq 0.618 \quad (12)$$

when all reaction and postreaction solvents and any catalysts used are recycled and that the reaction is run with no excess reagents. A plot of atom economy versus reaction yield shown in Figure 3a illustrates the domains of each parameter as well as the “green” zone defined by the upper right corner that satisfies inequality 12. One can think of the unit square as a probability domain space for all possible reaction outcomes run under stoichiometric conditions and complete solvent recovery and/or elimination with atom economies and reaction yields each falling between 0 and 1. The question being asked is what is the probability that the multiplicative product of “coordinates” (ϵ, AE) for a given reaction, which is its reaction mass efficiency, will have a value of at least 0.618. For a general threshold criterion α with a magnitude between 0 and 1, the probability that such a given reaction has an RME above α is given geometrically by the area above the curve $AE = (\alpha/\epsilon)$ and bounded by the limits $AE = 1$ and $\epsilon = 1$. Hence,

$$\begin{aligned} \text{probability} &= \frac{\text{area of green zone}}{\text{area of unit square}} \\ &= 1 - \alpha - \alpha \int_{\phi}^1 \frac{d\epsilon}{\epsilon} \quad (13) \\ &= 1 - \alpha + \alpha \ln \alpha \\ &= 0.0845 \end{aligned}$$

when $\alpha = \phi = 0.618$. Hence, the probability that a general chemical reaction run under stoichiometric conditions with complete recovery and/or elimination of reaction and postreaction solvents and catalysts meets the “golden” threshold criterion is only 8.45%! This value is in fact the minimum probability. If we consider a specific reaction and determine its minimum atom economy according to a Markush analysis previously discussed, then the magnitude of this probability is improved. There are three cases to consider, depending on whether the magnitude of $(AE)_{\min}$ is greater than, equal to, or less than the threshold parameter α . These are shown graphically in Figure 3b. Since the minimum atom economy represents a lower limit for AE, the allowable range of AE is $(AE)_{\min} \leq AE \leq 1$ and the resultant probability as a function of threshold is given by

probability =

$$\begin{cases} \frac{1 - \alpha + \alpha \ln \alpha}{1 - (AE)_{\min}}, & \text{if } (AE)_{\min} < \alpha \\ \frac{1 - \alpha + \alpha \ln \alpha}{1 - \alpha}, & \text{if } (AE)_{\min} = \alpha \\ \frac{1 - (AE)_{\min} + \alpha \ln(AE)_{\min}}{1 - (AE)_{\min}}, & \text{if } \alpha < (AE)_{\min} < 1 \end{cases} \quad (14)$$

The minimum probability is given by $1 - \alpha + \alpha \ln \alpha$ (equal to 0.0845 or 8.45% when $\alpha = 0.618$) when $(AE)_{\min} = 0$,

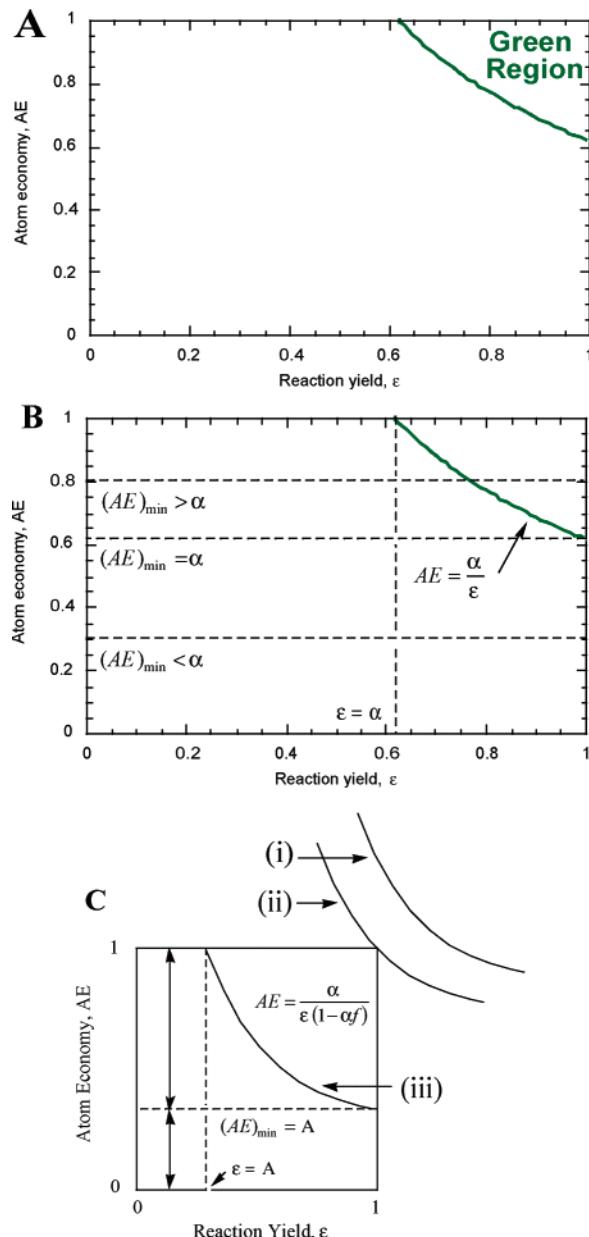


Figure 3. (A) Plot of atom economy versus reaction yield showing the domain space for stoichiometric chemical reactions satisfying the “golden” threshold criterion designated as the “green” zone under the assumption of complete reaction and postreaction solvent and catalyst recovery. (B) Same plot as in (A) showing the regions of $(AE)_{\min}$ relative to the threshold α . (C) Plot of atom economy versus reaction yield showing the domain space for stoichiometric chemical reactions satisfying the threshold criterion α under the assumption of no solvent and catalyst recovery: (i) $f > (1 - \alpha)/\alpha$, (ii) $f = (1 - \alpha)/\alpha$, and (iii) $f < (1 - \alpha)/\alpha$. Variable definitions are given in text.

and the maximum probability is given by $1 - \alpha$ (equal to 0.382 or 38.2% when $\alpha = 0.618$) when $(AE)_{\min} = 1$. As an example, for the oxidation of olefins to epoxides using *tert*-butylhydroperoxide $(AE)_{\min} = 0.37$, and the probability that this reaction will fall in the “green” zone is 13.4% when $\alpha = \phi = 0.618$. Similarly if hydrogen peroxide is used as the oxidizing agent, $(AE)_{\min} = 0.71$, and the corresponding probability is 27%. If a reaction has a minimum atom economy equal to the “golden” threshold, then the probability that the reaction mass efficiency is at least 0.618 is 22.1%.

If we consider nonstoichiometric reactions but maintain the constraint of reaction solvent recovery and/or elimination and catalyst recovery, then the threshold criterion for RME from eq 3 now becomes

$$\text{RME} = \epsilon(\text{AE}) \left[\frac{1}{\text{SF}} \right] \geq \alpha \quad (15)$$

If we let $\beta = (1/\text{SF})$ where $0 \leq \beta \leq 1$, then the probability of satisfying eq 15 is given geometrically by the intersecting volume above the hyperbolic surface $\text{AE} = (\alpha/\epsilon\beta)$ and bounded by the unit cube defined by $0 \leq \beta \leq 1$, $0 \leq \epsilon \leq 1$, and $0 \leq \text{AE} \leq 1$.⁶ Hence, we have

$$\begin{aligned} \text{probability} &= \frac{\text{volume of green zone}}{\text{volume of unit cube}} \\ &= 1 - \alpha + \alpha \ln \alpha - \frac{1}{2} \alpha (\ln \alpha)^2 \\ &= 0.0130 \end{aligned} \quad (16)$$

or 1.30% when $\alpha = \phi = 0.618$. Again, this represents a minimum probability estimate. When comparing this value with the previous one of 8.45%, this result confirms the expectation of a lower probability when excess reagents are designated as waste and is consistent with an attenuation of RME. We can see that avoiding excess reagents in reaction processes improves the likelihood of achieving the “golden” threshold criterion for reactions by about 85%, assuming that reaction and postreaction solvents are reclaimed and/or eliminated. When considering a given chemical reaction with a minimum atom economy, $(\text{AE})_{\min}$, we have by analogy with eq 14 three cases depending on the magnitude of $(\text{AE})_{\min}$ relative to the threshold α for the relevant probabilities

probability =

$$\begin{cases} \frac{1 - \alpha + \alpha \ln \alpha - 0.5\alpha(\ln \alpha)^2}{1 - (\text{AE})_{\min}}, & \text{if } (\text{AE})_{\min} < \alpha \\ \frac{1 - \alpha + \alpha \ln \alpha - 0.5\alpha(\ln \alpha)^2}{1 - \alpha}, & \text{if } (\text{AE})_{\min} = \alpha \\ \frac{1 - (\text{AE})_{\min} + \alpha \ln(\text{AE})_{\min} [1 + 0.5(\ln(\text{AE})_{\min} - 2 \ln \alpha)]}{1 - (\text{AE})_{\min}}, & \text{if } \alpha < (\text{AE})_{\min} < 1 \end{cases} \quad (17)$$

The minimum probability is given by $1 - \alpha + \alpha \ln \alpha - 0.5\alpha(\ln \alpha)^2$ (equal to 0.0130 or 1.30% when $\alpha = 0.618$) when $(\text{AE})_{\min} = 0$, and the maximum probability is given by $1 - \alpha + \alpha \ln \alpha$ (equal to 0.0845 or 8.45% when $\alpha = 0.618$) when $(\text{AE})_{\min} = 1$. Note the much narrower range of probabilities for achieving the “golden” threshold when excess reagents are permitted. Returning to the example of the epoxidation of olefins, the probability that this reaction will fall in the “green” zone under nonstoichiometric conditions and complete solvent recovery is 4.02% when *tert*-butylhydroperoxide is used as oxidizing agent and 4.39% when hydrogen peroxide is used. If a given reaction has

$(\text{AE})_{\min} = \alpha$, the probability that its reaction mass efficiency is at least 0.618 under nonstoichiometric reaction conditions and complete recovery of solvents and catalysts is 3.40%.

When considering stoichiometric reactions with no solvent and catalyst recovery we have

$$\text{RME} = \frac{\epsilon(\text{AE})}{1 + f\epsilon(\text{AE})} \geq \alpha \quad (18)$$

and

$$\text{AE} \geq \frac{\alpha}{\epsilon(1 - f\alpha)} \quad (19)$$

where the solvent and catalyst environmental impact factor is restricted to the range $0 < f < (1 - \alpha)/\alpha$ for the probability that the RME will be at least equal to α is positive. The graphical scenario is depicted in Figure 3c. When $f \geq (1 - \alpha)/\alpha$, then the function given in eq 19 does not intersect the unit square shown, and this means that the probability of achieving an RME above the threshold α is zero. The lower the threshold α is set, the greater is the allowable tolerance on the solvent and catalyst environmental impact factor f . One observes that $f = 0.618$ when $\alpha = 0.618$ and that $f \rightarrow \infty$ as $\alpha \rightarrow 0$. Hence, by analogy with eq 13

probability =

$$\begin{cases} 1 - \left(\frac{\alpha}{1 - \alpha f} \right) \left[1 - \ln \left(\frac{\alpha}{1 - \alpha f} \right) \right], & \text{if } 0 < f < (1 - \alpha)/\alpha \\ 0, & \text{if } f \geq (1 - \alpha)/\alpha \end{cases} \quad (20)$$

When $f = 0$, that is, when all solvents and catalysts are recovered, eq 20 reduces to eq 13 as expected. When considering a given reaction with a minimum atom economy of $(\text{AE})_{\min}$ we have

$$\text{probability} = \begin{cases} \frac{1 - A + A \ln A}{1 - (\text{AE})_{\min}}, & \text{if } (\text{AE})_{\min} < A \\ \frac{1 - A + A \ln A}{1 - A}, & \text{if } (\text{AE})_{\min} = A \\ \frac{1 - (\text{AE})_{\min} + A \ln(\text{AE})_{\min}}{1 - (\text{AE})_{\min}}, & \text{if } A < (\text{AE})_{\min} < 1 \end{cases} \quad (21)$$

where $0 < f < (1 - \alpha)/\alpha$ and the threshold for the atom economy $A = \alpha/(1 - \alpha f)$ applies. The probability that a given reaction achieves the threshold criterion $\text{RME} \geq \alpha$ falls in the range $1 - A + A \ln A < \text{probability} < 1 - A$. Typically, values of f are very much larger than 1 since solvents constitute the bulk mass of reaction ingredients. This effectively guarantees that the probability of achieving the requirement $\text{RME} \geq \alpha$ is zero regardless of the threshold α . The present analysis really points to the fact that to have a chance of meeting even a modest threshold of “greenness” it is imperative that solvents in reaction and postreaction operations be reclaimed, recycled, or altogether eliminated from processes. Again this result is consistent with the conclusion made by Eissen and Metzger that “reduction of solvent use is of great importance for the development of

(6) Edwards, C. H., Jr.; Penney, D. E. *Calculus and Analytical Geometry*, 3rd ed.: Prentice Hall: Englewood Cliffs, NJ, 1990; p 774 ff.

environmentally more benign processes".⁷

When considering reactions run under nonstoichiometric conditions with no solvent and catalyst recovery, the worst scenario in terms of waste production, we have by analogy with eq 20

probability =

$$\begin{cases} 1 - \left(\frac{\alpha}{1 - \alpha f} \right) \left[1 - \ln \left(\frac{\alpha}{1 - \alpha f} \right) + 0.5 \left[\ln \left(\frac{\alpha}{1 - \alpha f} \right) \right]^2 \right], \\ \quad \text{if } 0 < f < (1 - \alpha)/\alpha \\ 0, \text{ if } f \geq (1 - \alpha)/\alpha \end{cases} \quad (22)$$

and with eq 21

probability =

$$\begin{cases} \frac{1 - A + A \ln A - 0.5A(\ln A)^2}{1 - (AE)_{\min}}, \text{ if } (AE)_{\min} < A \\ \frac{1 - A + A \ln A - 0.5A(\ln A)^2}{1 - A}, \text{ if } (AE)_{\min} = A \\ \frac{1 - (AE)_{\min} + A \ln(AE)_{\min} [1 + 0.5(\ln(AE)_{\min} - 2 \ln A)]}{1 - (AE)_{\min}}, \\ \quad \text{if } A < (AE)_{\min} < 1 \end{cases}, \quad (23)$$

where $0 < f < (1 - \alpha)/\alpha$ and the threshold for the atom economy $A = \alpha/(1 - \alpha f)$ applies. Note that eqs 22 and 23 reduce to eqs 16 and 17 when $f = 0$.

All of the above analyses clearly demonstrate that chemical ingenuity in devising highly atom economical reactions from the outset is vital for increasing the odds of achieving the "golden" threshold. One therefore needs to be ahead of the threshold as early as possible at the kernel metrics level since other waste contributing factors such as excess reagents, solvents, and catalysts will necessarily rapidly diminish RME and amplify the Sheldon E-factor. For example, if a reaction run under stoichiometric conditions has a borderline AE of 0.618, then for it to have a RME of at least 0.618, its yield must be 100% and all reaction and postreaction solvents must be reclaimed and/or eliminated from the process. If the reaction has an AE of less than 0.618 from the outset then there is no hope that its RME will be at least 0.618 since the other factors will guarantee attenuation.

Application of the "golden" threshold criterion to a total synthesis of a complex target product by a linear sequence run under stoichiometric reaction conditions and complete solvent and catalyst recovery leads to the following inequality for the overall kernel environmental impact factor based on molecular weight,¹

$$E_{\text{mw}}^{\text{overall}} = \sum_{j=1}^n E_{\text{mw},j} \left(\frac{p_j}{p_n} \right) \leq 0.618 \quad (24)$$

where $E_{\text{mw},j}$ is the E-factor based on molecular weight for step j , p_j is the molecular weight of intermediate product formed in the j th step, and p_n is the molecular weight of the final product after n steps. If each step in the synthesis has

an E_{mw} factor of at most 0.618, then the best-case scenario for the overall E_{mw} factor is

$$(E_{\text{mw}}^{\text{overall}})_{\min} = 0.618 \sum_{j=1}^n \frac{p_j}{p_n} \quad (25)$$

It can be seen that for eq 25 to meet the "golden" threshold imposed by inequality 24 the summed term in eq 25 must be equal to 1 or that $\sum_{j=1}^n p_j = p_n$. This means that the "golden" threshold is achievable under the following two limiting conditions: (1) a one-step reaction where all required reagents are assembled at once or in a tandem fashion or (2) a sequence that involves continuous building from structure to structure toward the target product without producing any waste byproducts along the way and without isolations of intermediate products. Any other scenario guarantees that the overall E_{mw} factor will exceed 0.618 and hence that $(AE)_{\text{overall}}$ will be below 0.618. Moreover, the condition $\sum_{j=1}^n p_j = p_n$ holds for any threshold value α between 0 and 1 for $(E_{\text{mw}}^{\text{overall}})_{\min} = \alpha \sum_{j=1}^n (p_j/p_n) = \alpha$ to be true.

The mathematical arguments presented above are a clear impetus for the pursuit of total syntheses of complex targets either by single-step or tandem multicomponent reactions or reaction types that have atom economies of 100%. The new challenge in organic synthesis is to devise synthetic strategies with this as the central core paradigm of retrosynthetic plans. All of the ideas put forward here frame the task of inventing truly "green" individual reactions and "green" synthetic plans in accessible quantitative terms. It is evident that the vocabulary of organic transformations that make up the present library of reactions needs to be amplified to meet this challenge as will be shown in the next section which surveys the "green" performances of the standard repertoire of established organic reactions.

5. Patterns and Trends with Respect to Atom Economy for Various Reaction Types

Concepts presented in the preceding sections have established the importance of optimizing the related kernel metrics AE and E_{mw} . The focus is therefore on the intrinsic chemistry of chemical reactions that will really determine whether "green" reactions are possible. Therefore, we sought to implement these ideas to a set of named organic reactions that form the primary library of reactions that synthetic organic chemists draw upon in designing synthetic sequences to important target molecules.⁸ The Supporting Information contains tables of over 400 of these named organic reactions classified according to the following categories (Tables D1–D10): multicomponent reactions, carbon–carbon bond-

(8) (a) O'Neil, M. J., Smith, A., Heckelman, P. E., Eds. *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*, 13th ed.; Merck & Co., Inc.: Rahway, NJ, 2001. (b) Mundy, B. P., Ellerd, M. G. *Name Reactions and Reagents in Organic Synthesis*; Wiley: New York, 1988. (c) Laue, T.; Plagens, A. *Named Organic Reactions*; Wiley: New York, 2000. (d) Smith, J. G.; Fieser, M. *Fieser and Fieser's Reagents for Organic Synthesis Collective Index for Volumes 1–12*; Wiley: New York, 1990. (e) Li, J. J. *Name Reactions: A Collection of Detailed Reaction Mechanisms*; Springer-Verlag: Berlin, 2002. (f) Andraos, J. *Named Things in Chemistry and Physics*; <http://www.careerchem.com/NAMED/Homepage.html> (accessed July 2004).

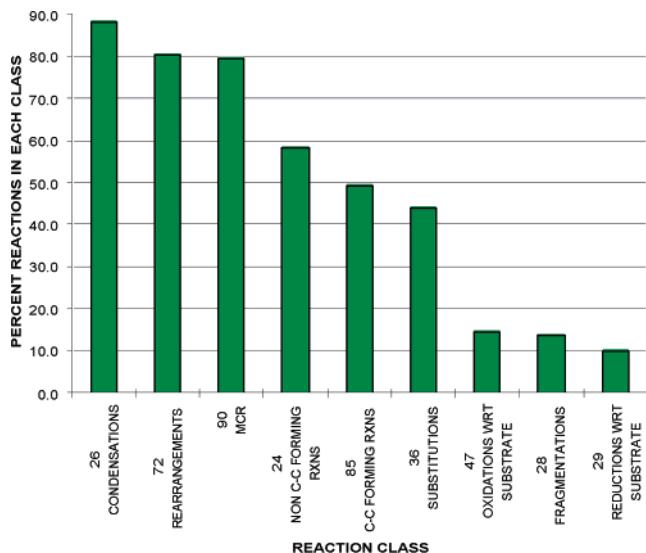


Figure 4. Histogram showing the proportion of “golden atom economical” reactions in each class of organic reaction. The numbers given in the abscissa correspond to the number of reactions in each reaction class considered in the database.

forming reactions, non-carbon–carbon bond-forming reactions, condensations, oxidations with respect to substrate, reductions with respect to substrate, rearrangements, substitutions, fragmentations and eliminations, and sequences. For each reaction a general balanced chemical equation showing all byproducts is given, based on appropriate Markush structures. From this equation estimates of AE_{\min} and E_{\max} are determined. Figure 4 presents the fraction of organic reactions in each category that meet the “golden” atom economy threshold and Figure 5 shows a plot of the cumulative growth of all named organic reactions discovered over time that are above or below it. For each reaction class Figure 6 shows a series of pairwise graphs that present the proportion of reactions above 60% according to the form of eq 1 given in Figure 2b and the rate of progress of their discovery since Wöhler’s urea synthesis in 1828 which is widely accepted as the dawn of organic synthesis. This analysis is instructive since it provides organic chemists with an opportunity to gauge the progress of reaction discovery

and to identify areas of research in organic reaction development that may need attention according to current interests in improving the “green” performance of organic transformations. The effect of representing the data in the form of Figure 2b is that data points satisfying the condition $AE > E_{\max}$ lie above the horizontal cutoff of $AE = 0.618$ and are conveniently spread out. The following section details trends among reaction classes according to this threshold criterion.

i. Multicomponent Reactions (MCRs).⁹ Although this class of reactions has a modern name, such reactions have been documented as early as 1850 with the discovery of the Strecker synthesis of α -cyanoamines as noted in Ugi’s historical review.^{9b} Other 19th century MCRs include the Hantzsch dihydropyridine synthesis (1882), Radziszewski imidazole synthesis reaction 1882, Riehm quinoline synthesis (1885), Doebner reaction 1887, Pinner triazine synthesis (1890), Hantzsch synthesis of pyrroles (1890), Biginelli reaction (1893), Guareschi–Thorpe condensation (1896), and Thiele reaction (1898). Of these the Biginelli reaction has received considerable attention in recent years.^{9f} It is interesting to note the two “bursts” of research activity in 1986 and again since 2003 where this class of reactions is now a rapidly developing field of organic synthesis. Reactions in this class are characterized by having at least three substrates reacting together in one reaction vessel either sequentially in a tandem or domino fashion or all at once. An important *caveat* in this kind of reaction is that the order of addition of substrates may matter in determining the final product of the reaction. It is obvious that the atom economy for the reaction increases as the molecular weights of the combining fragments increases. Reactions in this category which have $(AE)_{\min}$ values above the “golden” threshold are summarized in Table 4 along with relevant probabilities for achieving this criterion under stoichiometric reaction conditions with complete solvent and catalyst recovery.

ii. Condensations. Condensations are characterized by carbon–carbon couplings via straight-chain additions or cyclizations that produce as a byproduct some hydroxyllic molecule such as water or an alcohol. For this category as in the case of MCRs the atom economies increase as the

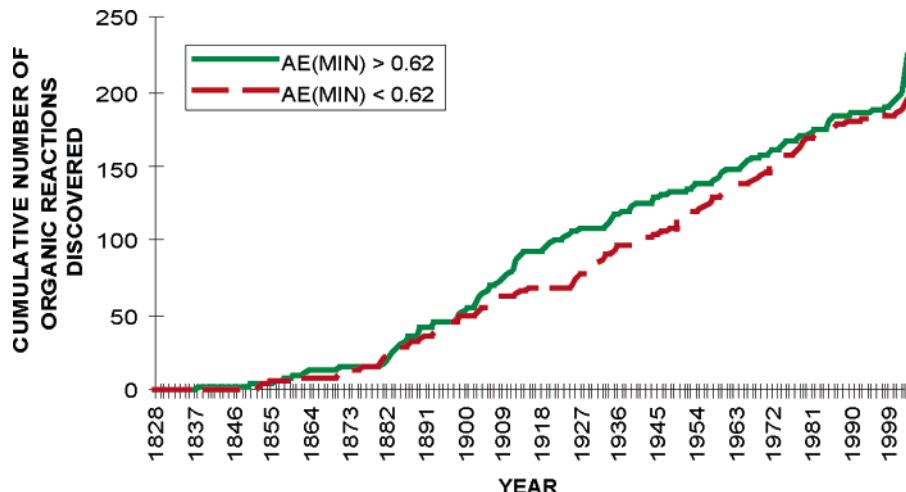
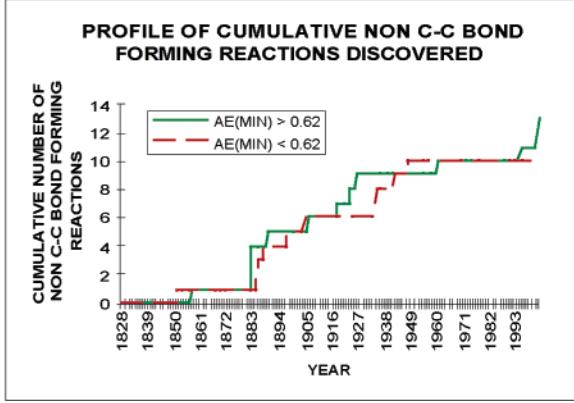
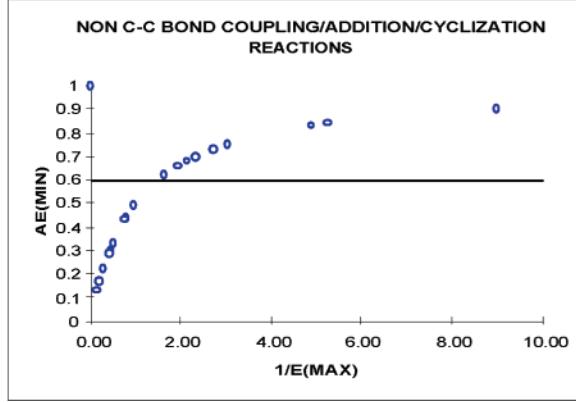
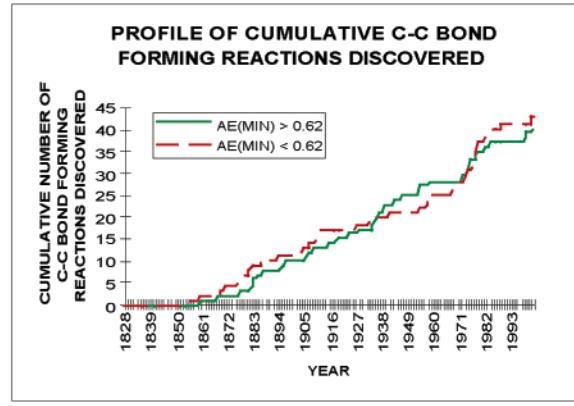
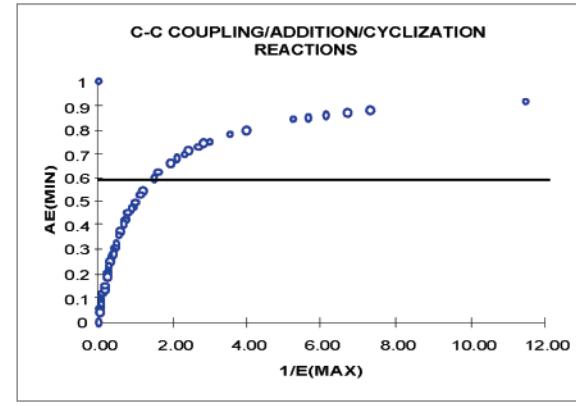
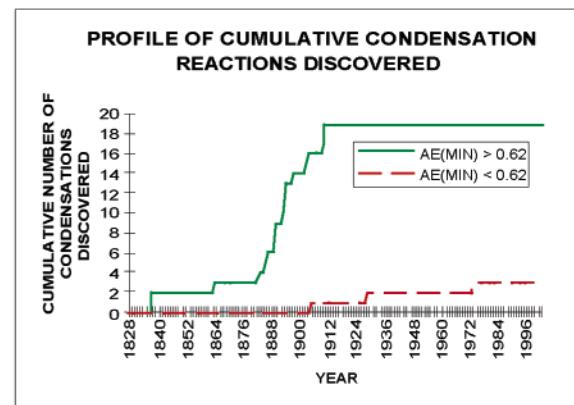
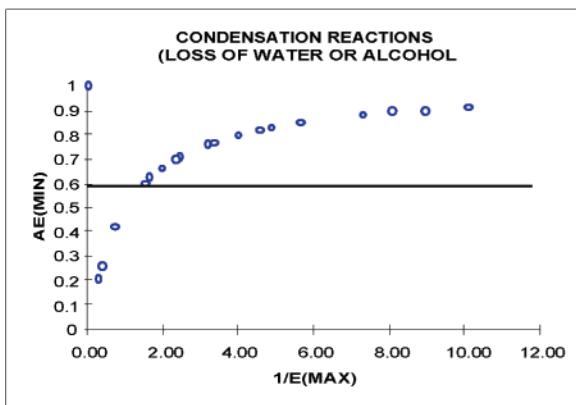
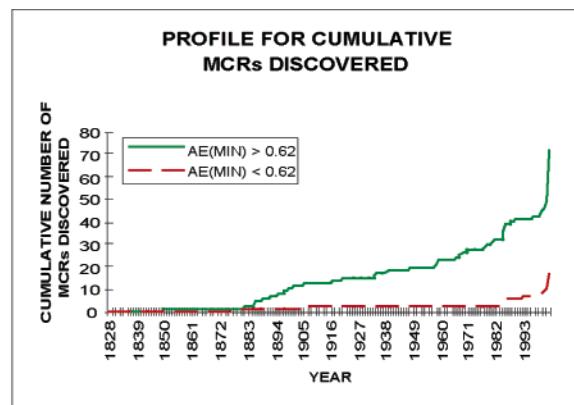
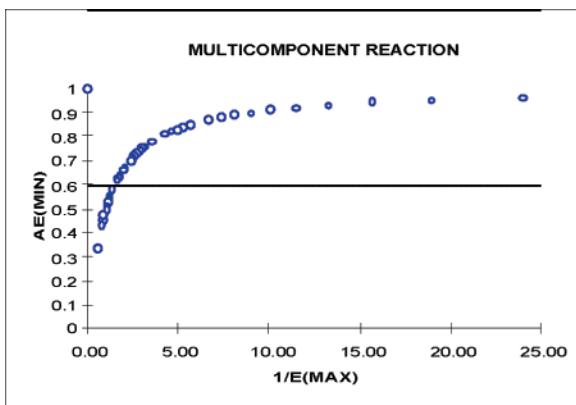
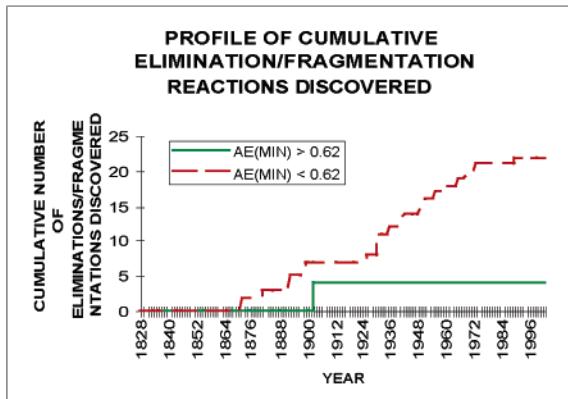
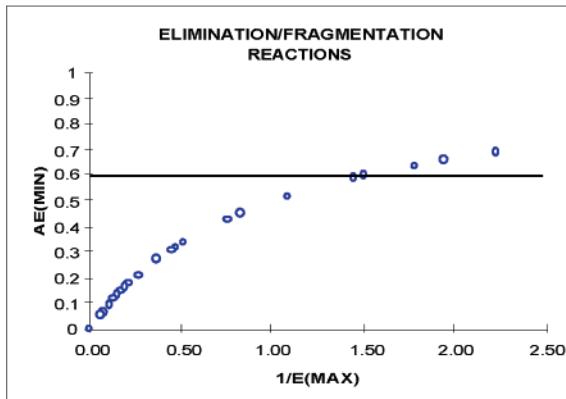
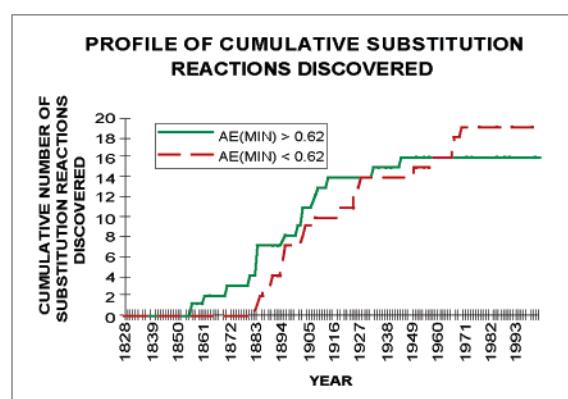
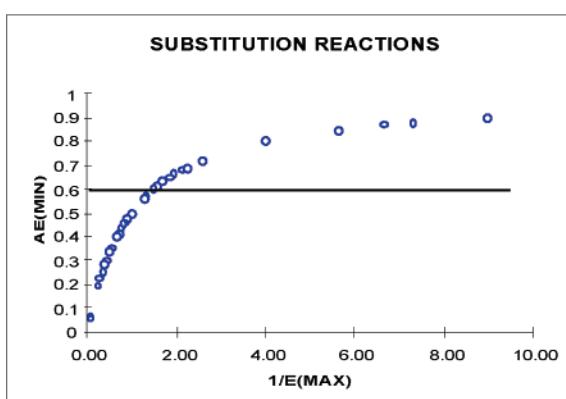
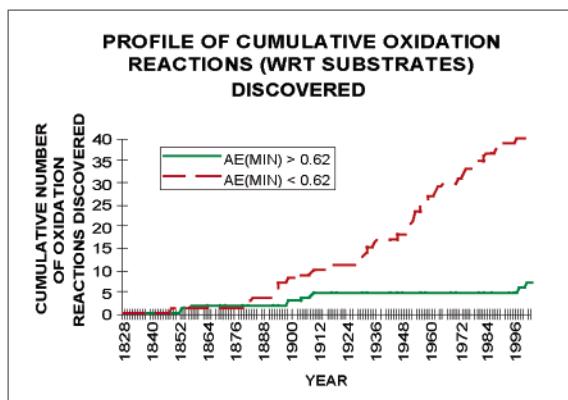
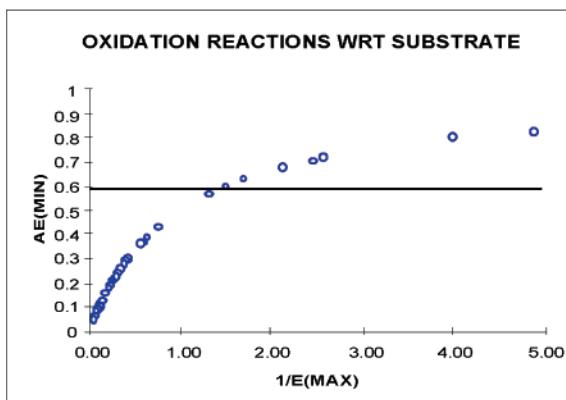
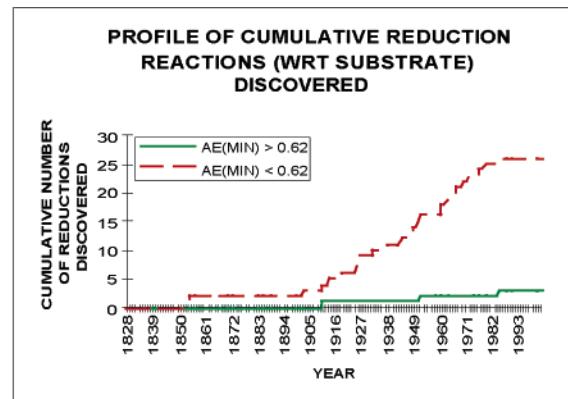
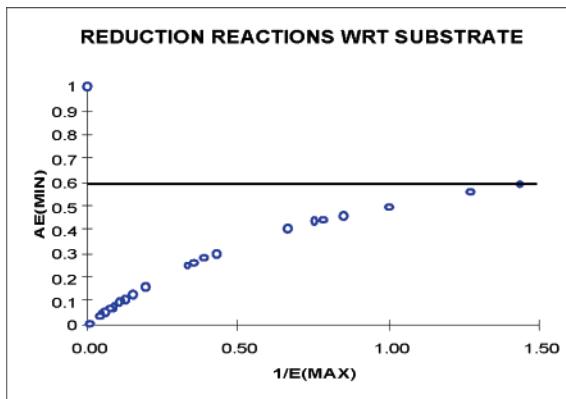


Figure 5. Plot showing the cumulative growth of organic reactions discovered that are above and below the “golden” atom economical threshold with the year of their discovery.





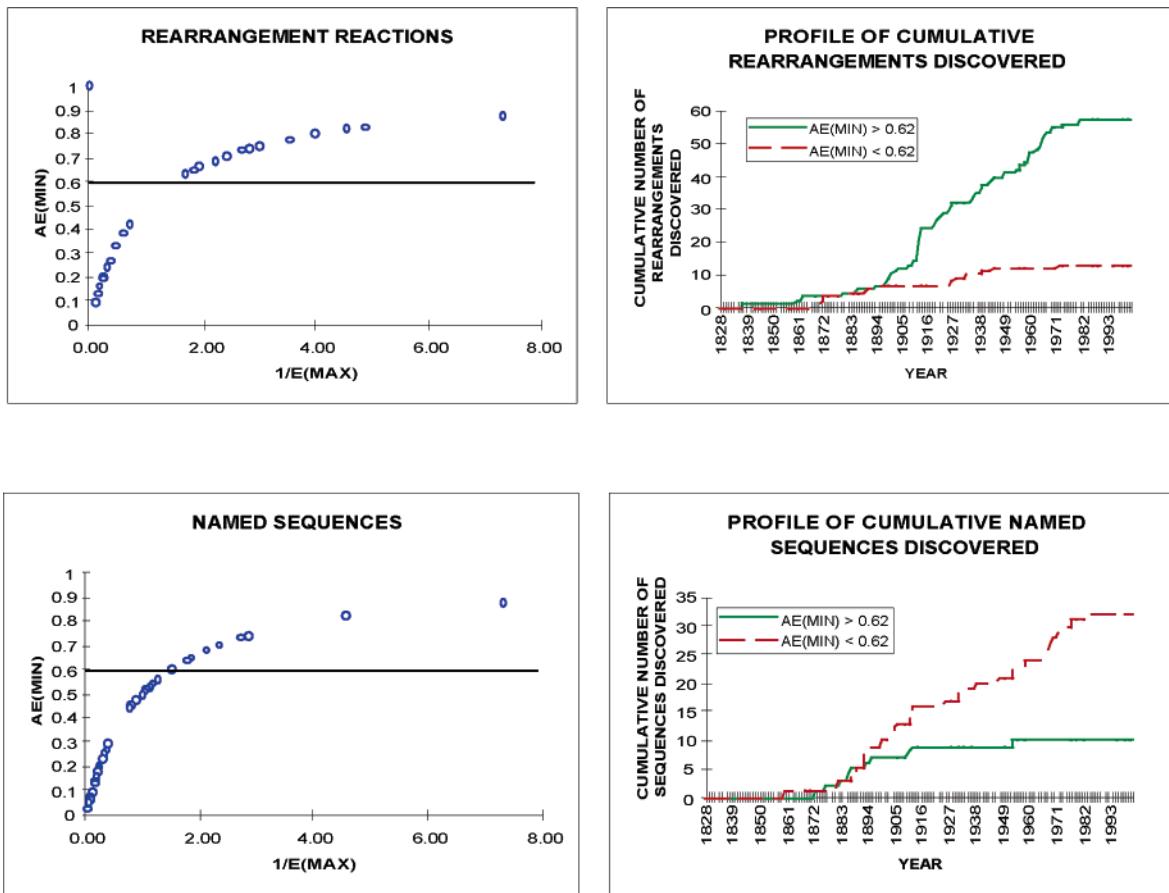


Figure 6. Pairwise plots showing proportion of reactions in each class that meet the “golden” atom economical threshold (points above horizontal line) and their discovery over time since Wöhler’s urea synthesis of 1828. Note that reactions with atom economies of 100% are arbitrarily plotted at the point (0, 1) so that they appear on the scale of the lefthand graphs.

molecular weights of the combining fragments increase. For cyclization reactions such as the Dieckmann condensation and the synthesis of cyclic ethers from straight-chain diols the atom economy increases with increasing ring size. About 90% of reactions in this class have minimum atom economies above the “golden” threshold. The exceptions are the Nenitzescu indole synthesis (1929), Mukaiyama aldol condensation (1974), and acyloin condensation (1905) which have $(AE)_{min}$ values of 0.42, 0.27, and 0.21, respectively.

iii. Carbon–Carbon and Non-carbon–Carbon Coupling Reactions. Again, these classes of reactions follow the pattern of other skeletal building reactions (MCRs and condensations) and are characterized by increasing atom economies as the molecular weights of the combining fragments increase. It is interesting to note that the progress of

discovery of reactions in this class with $(AE)_{min}$ values above the “golden” threshold is about the same as that with $(AE)_{min}$ values below it. Tables 5 and 6 summarize those reactions in each of these classes that meet the “golden” threshold.

iv. Reductions and Oxidations with Respect to Substrate. Redox-type reactions show the worst performance in meeting the “golden” atom economical threshold. Three reductions meet this criterion with $(AE)_{min}$ values of 1: hydrogenation of olefins using the Lindlar catalyst (1952), Noyori stereoselective hydrogenation reaction (1985), and Zincke disulfide cleavage reaction (1911), whereas, oxidations meeting this criterion are summarized in Table 7.

This class of reaction is particularly characterized by the production of significant waste byproducts which are the result of oxidation or reduction of reducing and oxidizing reagents, respectively. Recycling of byproducts back to the original oxidizing or reducing reagents necessarily involves at least another redox couple. It should be noted that the feasibility of recycling oxidizing and reducing reagents is strongly linked to energetics, that is, to appropriately chosen reagents with the right reduction potentials of half reactions so that the recycling process is made energetically favourable.¹⁰ Examples of the use of cascade redox reactions where the ultimate waste product is water have recently been developed.¹¹ The optimum oxidizing agents are molecular

(9) (a) Posner, G. H. *Chem. Rev.* **1986**, *86*, 831. (b) Ugi, I.; Dömling, A.; Hörl, W. *Endeavour* (new series) **1994**, *18*, 115. (c) Lombardo, M.; Trombini, C. *Semin. Org. Synth.* **1998**, *7*. (d) Ugi, I. *Proc. Est. Acad. Sci., Chem.* **1998**, *47*, 107. (e) Ugi, I.; Dömling, A.; Werner, B. *J. Heterocycl. Chem.* **2000**, *37*, 647. (f) Kappe, C. O. *Acc. Chem. Res.* **2000**, *33*, 879. (g) Dömling, A. *Curr. Opin. Chem. Biol.* **2000**, *4*, 318. (h) Ugi, I. *Pure Appl. Chem.* **2001**, *73*, 187. (i) Dömling, A. *Curr. Opin. Chem. Biol.* **2002**, *6*, 306. (j) Weber, L. *Drug Discovery Today* **2002**, *7*, 143. (k) Weber, L. *Curr. Med. Chem.* **2002**, *9*, 1241. (l) Orru, R. V. A.; de Greef, M. *Synthesis* **2003**, 1471. (m) von Wangelin, A.; Neumann, H.; Gördes, D.; Klaus, S.; Strübing, D.; Beller, M. *Chem. Eur. J.* **2003**, *9*, 4286. (n) Ugi, I.; Werner, B.; Dömling, A. *Molecules* **2003**, *8*, 53. (o) Simon, C.; Constantieux, T.; Rodriguez, J. *Eur. J. Org. Chem.* **2004**, 4957. (p) Ramachary, D. B.; Barbas, C. F., III *Chem. Eur. J.* **2004**, *10*, 5323.

Table 4. Multicomponent reactions discovered with $(AE)_{min}$ above the “golden” threshold

$(AE)_{min}$	p^a	name of reaction	year of discovery	$(AE)_{min}$	p^a	name of reaction	year of discovery
1	0.38	Ugi 3-CC MCR	1997	0.87	0.34	Petrenko-Kritschenko	1906
1	0.38	Alper carbonylation	1981	0.85	0.33	1,4-dihydropyridine synthesis	2003
1	0.38	β -iminoamine synthesis	2003	0.85	0.33	Bucherer hydantoin synthesis	1934
1	0.38	isoquinoline + DEAD + benzoquinones	2003	0.84	0.33	Ugi condensation	1959
				0.83	0.32	Mannich condensation	1912
1	0.38	Passerini	1921	0.83	0.32	oxazolidin-2-one synthesis	2003
1	0.38	Pauson-Khand	1971	0.83	0.32	[1 + 2 + 3] cyclohexanone synthesis (methyl ester)	1986
1	0.38	Reppe synthesis	1939				
1	0.38	Roelen	1948	0.82	0.32	fused 3-aminoimidazole synthesis	2003
1	0.38	Trost synthesis of 1,5-diketones	1997	0.82	0.32	1,5-benzodiazepine synthesis	2003
1	0.38	Trost's 4-CC γ, δ -unsaturated ketones	2000	0.82	0.32	Doebner	1887
1	0.38	[2 + 2 + 2] 1,6-aldo	1986	0.81	0.31	Hantzsch dihydropyridine synthesis	1882
1	0.38	[2 + 2 + 2] 1,6-Mannich	1975	0.81	0.31	3,4-dihydropyrimidin-2(1H)-one synthesis	2003
1	0.38	[2 + 2 + 2] alkene–alkyne annulation	1978	0.81	0.31	Rothemund	1935
1	0.38	[2 + 2 + 2] butadiene trimerization	1969	0.81	0.31	Thiele	1898
1	0.38	[2 + 2 + 2] trimerization of alkynes	1985	0.78	0.30	1H-pyrrolo[3,2-e]-1,2,4-triazine synthesis	2003
1	0.38	[2 + 2 + 2] trimerization of arylisocyanates	1985	0.76	0.29	α -amino acids via münchnone synthesis	2003
1	0.38	fused benzochromene synthesis	2003	0.76	0.29	Radziszewski-type microwave irradiation reaction	2003
1	0.38	β -acetoamido carbonyl compounds	2003	0.75	0.29	Strecker synthesis of α -cyanoamines	1850
0.96	0.37	Knoevenagel hetero-Diels–Alder reaction	2002	0.75	0.29	Riehm quinoline synthesis	1885
0.95	0.37	2',3'-pyranone (pyrrolidinone) fused tryptamine synthesis	2003	0.74	0.28	aldehyde + alkyne + oxadiazoline	2001
0.95	0.37	aminoalkylation of naphthols with chiral amines	2003	0.73	0.28	alkynes + imines + organoboron reagents	2003
0.95	0.37	Betti	1900	0.73	0.28	Grigg allenylation	2003
0.94	0.36	cis-isoquinolonic acid synthesis	2003	0.73	0.28	Biginelli synthesis	1893
0.94	0.36	pyranoquinoline synthesis	2003	0.73	0.28	Pinner triazine synthesis	1890
0.94	0.36	Oikawa (aldehyde + Meldrum's acid + indole)	1978	0.72	0.27	asymmetric Mannich-type reaction	2003
0.93	0.36	List (aldehyde + Meldrum's acid + ketone)	2001	0.72	0.27	[(1 + 2) + (1 + 2)] cyclopropanation (X = CH ₂)	1985
0.93	0.36	aza-Diels–Alder synthesis of tetrahydroquinolines	2003	0.71	0.27	[1 + (1 + 2)] cyclopropanation	1985
0.93	0.36	bis(indoyl)methanes	2003	0.7	0.27	Asinger	1958
0.92	0.36	Tietze (aldehyde + Meldrum's acid + vinyl ether)	1990	0.67	0.25	Hantzsch synthesis of pyrroles	1890
0.91	0.35	Grieco condensation	1988	0.66	0.24	Nenitzescu synthesis of pyrylium salts (perchlorate salt, acid chloride)	1959
0.91	0.35	Weiss	1968				
0.9	0.35	malonylurea + aldehyde + isocyanide	2002	0.66	0.24	Feist–Benary synthesis of pyrroles	1902
0.9	0.35	[2 + 2 + 2] isocyanate–ketone cyclization	1986	0.64	0.23	amide + thioisocyanate + alkyl chloride	2002
0.89	0.35	Gewald aminothiophene synthesis	1966	0.64	0.23	Guareschi–Thorpe condensation (R = Me)	1896
0.88	0.34	Robinson–Scheopf synthesis of tropane	1917	0.64	0.23	indolizine synthesis	2003
0.87	0.34	[2 + 2 + 2] vinylphosphonium bromide–ketone cyclization	1981	0.63	0.23	Dornow–Wiehler isoxazole synthesis	1952
				0.62	0.22	tandem Petasis–Ugi condensation	2003

^a Probability that RME meets the “golden” threshold under stoichiometric reaction conditions and complete solvent and catalyst recovery.

oxygen from air and hydrogen peroxide which both produce water as a byproduct. Noyori and others have recently developed “green” catalytic oxidations using aqueous hy-

(10) (a) Noyori, R.; Aoki, M.; Sato, K. *Chem. Commun.* **2003**, 1977. (b) Sato, K.; Aoki, M.; Noyori, R. *Science* **1998**, 281, 1646. (c) Sheldon, R. A.; Arends, I. W. C. E.; Brink, G. J. T.; Dijksman, A. *Acc. Chem. Res.* **2002**, 35, 774.
 (11) (a) Latimer, W. M. *Oxidation Potentials*; Prentice-Hall: New Jersey, 1952. Dobos, D. *Electrochemical Data: A Handbook for Electrochemists in Industry and Universities*; Elsevier: Amsterdam, 1975. (b) Wardman, P. *J. Phys. Chem. Ref. Data* **1989**, 18, 1637. (c) Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*; Pergamon Press: Oxford, 1966. (d) Bard, A. J.; Parson, R.; Jordan, J. *Standard Potentials in Aqueous Solutions*; Marcel Dekker: New York, 1985. (e) Milazzo, G.; Caroli, S. *Tables of Standard Electrode Potentials*; Wiley: New York, 1978.
 (12) (a) Choudary, B. M.; Chowdari, N. S.; Mahdi, S.; Kantam, M. L. J. *Org. Chem.* **2003**, 68, 1736. (b) Kim, S. S.; Rajagopal, G. *Synth. Commun.* **2004**, 33, 2237. (c) Chandrasekhar, S.; Narasimulu, C.; Sultana, S. S.; Reddy, N. R. *Chem. Commun.* **2003**, 1716.

drogen peroxide and other oxidants.¹² On the other hand, the optimum reducing agent is hydrogen gas which produces no byproducts. The main objective of current research is to find suitable catalysts to effect redox reactions using any one of the optimum redox reagents. In this regard the Lindlar and Noyori hydrogenation reactions are good examples where this has been achieved. As seen by the graphs in Figure 5 progress in developing “green” redox reactions has been slow but is now gaining attention.

v. Substitution Reactions. Substitutions have the unique property that the atom economy increases if the in-coming group is heavier than the leaving group, otherwise it will decrease. The *caveat* is that good leaving groups tend to be large, as in the case of tosylate and iodide. It is this class of reactions that is employed in protection–deprotection strate-

Table 5. Carbon–carbon bond-forming reactions discovered with $(AE)_{min}$ above the “golden” threshold

$(AE)_{min}$	p^a	name of reaction	year of discovery	$(AE)_{min}$	p^a	name of reaction	year of discovery
1	0.38	Diels–Alder	1928	0.88	0.34	Houben–Hoesch synthesis	1915
1	0.38	Koch–Haaf carbonylation	1958	0.88	0.34	Jacobs oxidative coupling	2003
1	0.38	Michael 1,4-addition	1887	0.87	0.34	Fischer indole synthesis	1883
1	0.38	Danishefsky	1974	0.86	0.33	Meerwein arylation	1939
1	0.38	Kolbe synthesis	1860	0.85	0.33	Zincke–Suhl	1906
1	0.38	Paterno–Büchi	1909	0.84	0.33	Robinson annulation	1935
1	0.38	Gattermann–Koch	1897	0.8	0.31	Henry reaction (under strong acid conditions)	1895
1	0.38	Nazarov cyclization	1946			Friedlander synthesis	1882
1	0.38	Henry reaction (under weak acid conditions)	1895	0.78	0.30	Sonogashira (X = Cl)	1975
1	0.38	Bergmann cyclization	1972	0.74	0.28	Kilian–Fischer synthesis	1885
1	0.38	Tishchenko	1906	0.74	0.28	Eglinton	1954
1	0.38	Prins	1919	0.71	0.27	Hammick	1937
1	0.38	Danheiser’s alkyne–cyclobutanone cyclization	1984	0.71	0.27	Perkin	1868
1	0.38	Danheiser’s [4 + 4] annulation	1982	0.68	0.26	Marshallk Friedel–Crafts alkylation (phenyl, Cl case)	1939
1	0.38	Trost [5 + 2] annulation	2000			Suzuki coupling (Z = Cl)	1877
1	0.38	Trost allene–alkene addition	1999	0.66	0.24	Vorbrueggen coupling	1979
1	0.38	Baylis–Hillmann	1972	0.62	0.22	Gomberg–Bachmann reaction (biphenyl case)	1978
1	0.38	Nieuwland enyne synthesis	1934	0.6	0.21	cyclopropanation of olefins with diazomethane	1924
1	0.38	[1,5]-ene reaction	1943				
0.92	0.36	Doetz	1975	0.6	0.21		1935

^a Probability that RME meets the “golden” threshold under stoichiometric reaction conditions and complete solvent and catalyst recovery.

Table 6. Non-carbon–carbon bond-forming reactions discovered with $(AE)_{min}$ above the “golden” threshold

$(AE)_{min}$	p^a	name of reaction	year of discovery	$(AE)_{min}$	p^a	name of reaction	year of discovery
1	0.38	formation of hemiacetals or hemiketals	1907	0.73	0.28	formation of acetals or ketals	1924
1	0.38	Menshutkin	1890	0.73	0.28	Hoffmann–Löffler–Freytag (ring size = 7)	1883
1	0.38	1,3-dipolar cycloadditions	1961	0.7	0.27	Hofmann–Löffler–Freytag (ring size = 6)	1883
0.9	0.35	Stoltz aerobic etherification	2003	0.68	0.26	Buchwald–Hartwig cross coupling	1995
0.84	0.33	Stahl aerobic amination	2002			(X = Br, Y = Ph)	
0.83	0.32	azo coupling using diazonium chlorides	1858	0.66	0.24	Hofmann–Löffler–Freytag (ring size = 5)	1883
0.75	0.29	Polonovski reaction (nonrearranged product)	1927	0.62	0.22	Staudinger	1919

^a Probability that RME meets the “golden” threshold under stoichiometric reaction conditions and complete solvent and catalyst recovery.

Table 7. Oxidation reactions discovered with $(AE)_{min}$ above the “golden” threshold

$(AE)_{min}$	p^a	name of reaction	year of discovery
0.83	0.32	Harries ozonolysis of olefins with oxidative workup	1905
0.8	0.31	Willgerodt	1909
0.72	0.27	Baeyer–Villiger reaction with hydrogen peroxide	1899
0.71	0.27	epoxidation of olefins with hydrogen peroxide	2001
0.68	0.26	Cannizzaro reaction with respect to carboxylic acid product as target	1853
0.63	0.23	Uemura oxidation	1998
0.6	0.21	Griess diazotization reaction of anilines	1858

^a Probability that RME meets the “golden” threshold under stoichiometric reaction conditions and complete solvent and catalyst recovery.

gies in organic synthesis, and it is not surprising, therefore, that given the behaviour of AE with leaving group size minimization of this kind of strategy in synthetic sequences was named as one of the 12 Principles of Green Chemistry.¹³

(13) (a) Anastas, P. T.; Warner, J. C. *Green Chemistry: Theory and Practice*; Oxford University Press: New York, 1998. (b) Anastas, P. T.; Kirchhoff, M. M. *Acc. Chem. Res.* **2002**, 35, 686. (c) Anastas, P. T. *Green Chem.* **2003**, 5, G29. (d) Sheldon, R. A. *Top. Curr. Chem.* **1993**, 164, 21.

Table 8 lists named substitution reactions with $(AE)_{min}$ meeting the “golden” threshold.

vi. Fragmentation and Elimination Reactions. For fragmentation or elimination reactions it is expected that the proportion of high atom economical reactions is low since these reactions are the reverse of skeletal building-up reactions. The atom economy decreases as the molecular weight of the leaving fragment increases. The Tiffeneau–

Table 8. Substitution reactions discovered with $(AE)_{min}$ above the “golden” threshold

$(AE)_{min}$	p^a	name of reaction	year of discovery	$(AE)_{min}$	p^a	name of reaction	year of discovery
0.9	0.35	aromatic sulfonation	1871	0.69	0.26	Finkelstein reaction (exchange Cl for Br)	1910
0.9	0.35	Sanger	1945	0.69	0.26	Chichibabin	1914
0.9	0.35	Helferich (6 carbon sugars)	1933	0.68	0.26	esterification with diazomethane	1908
0.88	0.34	Sandmeyer (phenyl, I case)	1884	0.68	0.26	Lapworth reaction using chlorine	1903
0.87	0.34	aromatic nitration	1856	0.66	0.24	aromatic bromination	1862
0.85	0.33	Sandmeyer (phenyl, Br case)	1884	0.63	0.23	Hell–Volhard–Zelinsky	1881
0.8	0.31	Sandmeyer (phenyl, Cl case)	1884	0.61	0.22	Koenigs–Knorr synthesis	1901
0.72	0.27	Fischer esterification	1895	0.6	0.21	Lapworth reaction using bromine	1903

^a Probability that RME meets the “golden” threshold under stoichiometric reaction conditions and complete solvent and catalyst recovery.

Table 9. Rearrangement reactions involving rearrangements of intermediates discovered with $(AE)_{min}$ above the “golden” threshold

$(AE)_{min}$	p^a	name of reaction	year of discovery
0.88	0.34	Wagner–Meerwein	1899
0.83	0.32	acyl rearrangement ($X = Cl$)	1920
0.82	0.32	Wolff rearrangement to make cyclic methyl esters (ring size = 7)	1912
0.80	0.31	Wolff rearrangement to make cyclic methyl esters (ring size = 6)	1912
0.80	0.31	vinylogous Wolff	1974
0.78	0.30	Wolff rearrangement to make cyclic methyl esters (ring size = 5)	1912
0.75	0.29	Wolff rearrangement to make cyclic methyl esters (ring size = 4)	1912
0.74	0.28	Wittig [1,2]	1942
0.73	0.28	Favorskii rearrangement with substrate ring size = 7	1913
0.71	0.27	Favorskii rearrangement with substrate ring size = 6	1913
0.71	0.27	pinacol	1860
0.69	0.26	Favorskii rearrangement with substrate ring size = 5	1913
0.69	0.26	hydroboration–borane	1956
0.66	0.24	Favorskii rearrangement with substrate ring size = 4	1913
0.65	0.24	Nametkin ($X = Cl$)	1923
0.63	0.23	hydroperoxide rearrangement with respect to ketone as target product	1944

^a Probability that RME meets the “golden” threshold under stoichiometric reaction conditions and complete solvent and catalyst recovery.

Demjanov reaction is the only one in this category with $(AE)_{min}$ values meeting the “golden” threshold. The value of $(AE)_{min}$ increases with ring size: 0.69 for ring size 8, 0.66 for ring size 7, 0.64 for ring size 6, and 0.60 for ring size 5.

vii. Rearrangement Reactions. Rearrangements of substrates always have atom economies of 100%; however, about 41% of reactions in this class have atom economies less than unity. This is because they involve rearrangements of intermediates along the reaction pathway yet they are still categorized as rearrangement reactions. Examples of this type of rearrangement that meets the “golden” atom economical threshold are given in Table 9. For the Wolff and Favorskii rearrangements the atom economy increases with ring size as expected.

viii. Sequences. Only 24% of the 42 named sequences have overall minimum atom economies that meet the “golden” threshold. These are summarized in Table 10.

6. Recycling of Byproducts Outlook

From the database of chemically balanced named organic reactions presented in the previous section it is possible to determine the proportion in each class which either produce no byproducts, produce some byproducts that could be recycled to one of the starting materials, or produce byproducts which are not worth recycling. These are displayed graphically in Figure 7.

Table 10. Sequences discovered with $(AE)_{min}$ above the “golden” threshold

$(AE)_{min}$	p^a	name of reaction	year of discovery
0.88	0.34	Pictet–Spengler isoquinoline synthesis	1911
0.82	0.32	Doebner	1887
0.82	0.32	Pinner synthesis of guanidines	1877
0.74	0.28	Stork enamine reaction ($X = Cl$)	1954
0.73	0.28	Doebner–Miller	1894
0.70	0.27	von Richter cinnoline synthesis	1883
0.68	0.26	Pfitzinger	1886
0.65	0.24	Stolle synthesis of indoles	1913
0.64	0.23	Urech synthesis of hydantoins	1873
0.60	0.21	Pschorr	1896

^a Probability that RME meets the “golden” threshold under stoichiometric reaction conditions and complete solvent and catalyst recovery.

From this graph one can see a direct correlation between those reaction classes with low minimum atom economies below the “golden” threshold and those that have a high potential for recycling byproducts back to reagents. The best examples are the redox reactions which have the highest potential for recycling of byproducts (yellow bars) and the lowest proportion of “golden” reactions (green bars) (see Figure 7). Condensations, on the other hand, have the lowest proportional yellow bar and the highest proportional green bar. Tables 11 and 12 summarize the net gains or losses in

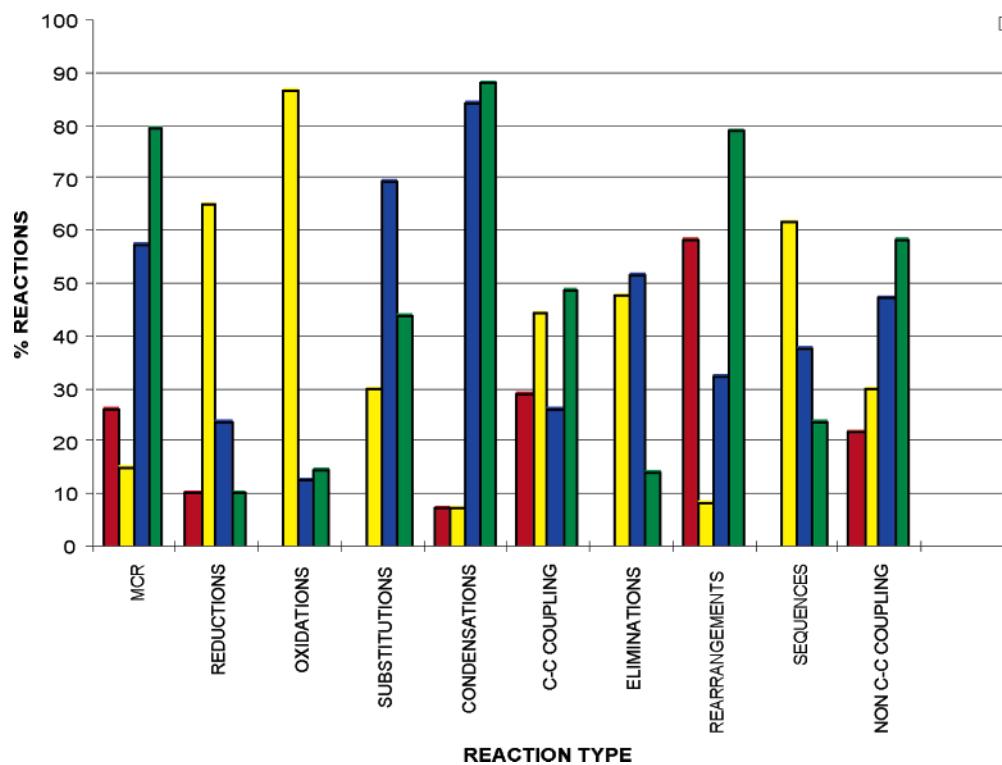
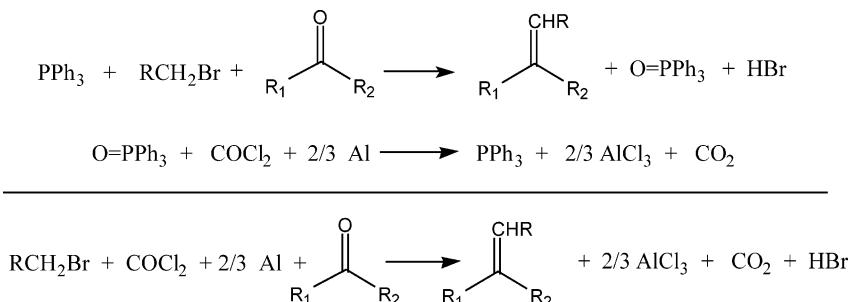


Figure 7. Recycling profiles for named organic reactions arranged by reaction class: (red) proportion of reactions producing no byproducts; (yellow) proportion of reactions where recycling of byproducts may be viable; (blue) proportion of reactions where recycling of byproducts is not likely viable; and (green) proportion of reactions that meet “golden” atom economical threshold.

Scheme 5



overall minimum atom economy for each named organic reaction and sequence if the original reaction(s) is(are) combined with a viable byproduct-recycling reaction under stoichiometric conditions. Tables R1 and R2 in the Supporting Information summarize literature reactions with yields of at least 90% that can be used to recycle waste byproducts back to reagents. The determinations in Tables 11 and 12 were made under the assumption that both the forward and recycling reactions have yields of 100% and that all the byproduct collected in the forward reaction is converted back to the appropriate starting reagent. Hence, the net gains in $(\text{AE})_{\min}$ reported represent upper limits and are therefore best-case scenarios. Inclusion of reaction yields for both the forward and recycling reactions will necessarily change these scenarios; however, the results of the present exercise are good enough to at least indicate which reactions have the greatest potential for increasing the overall atom economy when byproducts are recycled. The Wittig reaction is used

as an example illustrating the numerical calculation of the net gain in $(\text{AE})_{\min}$. Scheme 5 shows balanced chemical equations for the Wittig reaction, a potential recycling reaction for the conversion of triphenylphosphine oxide back to triphenylphosphine,¹⁴ and the sum of the two chemical equations.

The expressions for the environmental impact factor for the Wittig reaction and the combined Wittig and recycling reactions are given by eqs 26 and 27, respectively

$$E_{\text{mw}} = \frac{361.9}{x + 25} \quad (26)$$

$$E_{\text{mw}} = \frac{213.8}{x + 25} \quad (27)$$

where x is the sum of molecular weights of all the R groups.

(14) (a) Hermling, D.; Bassler, P.; Hammes, G.; 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.

Table 11. Summary of overall atom economy gains and losses for combined named organic and byproduct recycling reactions under the assumption of 100% yields for all reactions under stoichiometric conditions

% change in $(AE)_{min}$	$(AE)_{min}$ with recycling byproducts	$(AE)_{min}$ without recycling byproducts	named reaction
+1823.1	1	0.052	Nozaki reaction (X = Br)
+1122.2	0.44	0.036	McMurry reaction
+987.0	1	0.092	Nozaki reaction (X = Cl)
+957.1	0.74	0.07	reduction of nitroaromatics with iron (Bechamp reaction)
+803.8	0.47	0.052	Clemmensen reduction
+516.2	0.61	0.099	Burgess dehydration
+425.0	0.63	0.12	Criegee glycol cleavage reaction
+350.0	0.45	0.1	Sarett procedure
+305.3	0.77	0.19	Sharpless oxyamination ^a
+273.7	0.71	0.19	Shi asymmetric epoxidation
+263.6	0.8	0.22	synthesis of acetates via oxymercuration ^b
+260.0	0.72	0.2	Fehling reaction
+250.0	0.63	0.18	Lemieux–Johnson cleavage of olefins to aldehydes ^c
+150.0	0.3	0.12	oxymercuration of olefins
+248.8	0.15	0.043	thioketal desulfurization
+207.1	0.86	0.28	Elbs reaction
+200.0	0.63	0.21	Oppenauer oxidation ^d
+194.1	1	0.34	epoxide rearrangement
+189.9	0.2	0.069	Kochi reaction
+189.3	0.81	0.28	Wessely oxidation
+177.8	1	0.36	Lemieux–Johnson oxidation of olefins to 1,2-diols ^e
+163.3	0.79	0.3	Paal–Knorr synthesis of furans
+156.4	1	0.39	permanganate oxidation of olefins to 1,2-diols
+156.3	0.41	0.16	Glaser coupling reaction
+142.4	0.16	0.066	Dess–Martin oxidation
+140.5	0.89	0.37	Gewald aminothiophene synthesis
+137.5	0.19	0.08	Prevost reaction
+136.0	0.59	0.25	Skraup reaction
+131.0	0.67	0.29	Fenton reaction
+130.8	0.21	0.091	Corey–Kim reaction
+125.4	0.16	0.071	Corey–Bakshi–Shibata reduction
+125.0	0.27	0.12	Ruff–Fenton reaction
+122.2	1	0.45	tandem Passerini–Wittig reaction
+108.3	0.75	0.36	organocuprate Michael addition (R = Me)
+105.4	0.76	0.37	Riley oxidation
+100.0	0.19	0.095	Woodward <i>cis</i> -hydroxylation
+100.0	0.6	0.3	Hooker oxidation
+100.0	0.4	0.2	Hosomi–Sakurai reaction α,β -unsaturated ketones
+100.0	0.4	0.2	Mukaiyama–Michael reaction
+100.0	0.38	0.19	Sharpless oxyamination ^f
+95.7	0.45	0.23	Jones oxidation of secondary alcohols to ketones
+92.3	1	0.52	pinacol reaction
+92.3	0.25	0.13	Sharpless–Jacobsen dihydroxylation ^g
+91.9	0.71	0.37	Sharpless epoxidation
+90.5	0.4	0.21	Oppenauer oxidation ^h
+90.0	0.38	0.2	Wacker reaction
+85.4	0.76	0.41	Vilsmeier–Haack–Arnold reaction
+83.3	0.33	0.18	Borodin–Hunsdiecker reaction
+79.1	0.77	0.43	Harries ozonolysis of olefins with reductive workup ⁱ
+77.8	0.48	0.27	Mukaiyama aldol condensation
+76.8	0.0099	0.0056	radical dehalogenation using tributyltin hydride ^j
+72.1	0.74	0.43	Bucherer reaction
+70.0	0.51	0.3	Wöhler urea synthesis
+65.1	0.71	0.43	Jacobsen's epoxidation
+64.3	0.0092	0.0056	radical dehalogenation using tributyltin hydride ^k
+60.0	0.16	0.1	Midland reduction
+56.3	0.15	0.096	Pfitzner–Moffatt oxidation of primary alcohols to aldehydes ^l
+55.8	0.12	0.077	Corey–Winter reaction
+52.6	0.29	0.19	Cadiot–Chodkiewitz reaction
+52.5	0.61	0.4	Wohl–Ziegler bromination
+50.0	0.57	0.38	Reformatskii reaction
+48.0	0.37	0.25	Gribble reduction of diaryl ketones
+47.7	0.13	0.088	Swern oxidation
+47.3	0.81	0.55	Petasis condensation
+46.2	0.19	0.13	Heck reaction (X = Br)
+45.8	0.14	0.096	Pfitzner–Moffatt oxidation of primary alcohols to aldehydes ^m
+45.2	0.9	0.62	Vorbrueggen coupling

Table 11 (Continued)

% change in $(AE)_{min}$	$(AE)_{min}$ with recycling byproducts	$(AE)_{min}$ without recycling byproducts	named reaction
+43.9	0.59	0.41	Dakin–West reaction
+40.9	0.31	0.22	Wenker synthesis
+40.4	0.66	0.47	alkyne–carbamate–phosphine annulation ⁿ
+38.3	0.65	0.47	alkyne–carbamate–phosphine annulation ^o
+37.5	0.22	0.16	Corey–Chaykovsky epoxidation
+36.7	0.82	0.6	Cannizzaro reaction with respect to carboxylic acid product
+36.0	0.68	0.5	Schiemann reaction
+35.0	0.54	0.4	Cannizzaro reaction with respect to alcohol product
+34.6	0.35	0.26	Etard reaction
+30.2	0.56	0.43	Graham reaction (X = Br)
+29.2	0.31	0.24	Tiemann rearrangement ^p
+28.3	0.68	0.53	Ullmann coupling (biphenyl case, X = Cl)
+27.8	0.23	0.18	Lemieux–Johnson cleavage of olefins to aldehydes ^q
+25.8	0.78	0.62	tandem Petasis–Ugi condensation
+25.5	0.59	0.47	alkyne–carbamate–phosphine annulation ^r
+25.4	0.74	0.59	reduction of nitroaromatics with hydrogen sulfide
+25.0	0.11	0.088	Tebbe olefination
+25.0	0.12	0.096	Pfitzner–Moffatt oxidation of primary alcohols to aldehydes ^s
+24.2	0.41	0.33	[2 + 2 + 2] Michael–Michael–1,6-Wittig ^t
+23.6	0.068	0.055	Gabriel synthesis
+23.1	0.96	0.78	1 <i>H</i> -pyrrolo[3,2- <i>e</i>]-1,2,4-triazine synthesis
+21.4	0.34	0.28	Stephen reduction
+21.3	0.91	0.75	Polonovski reaction (nonrearranged product)
+21.2	0.4	0.33	[2 + 2 + 2] Michael–Michael–1,6-Wittig ^u
+20.0	0.24	0.2	Ramberg–Backlund rearrangement
+19.8	0.97	0.81	Thiele reaction
+19.0	0.25	0.21	Oppenauer oxidation ^v
+18.0	0.72	0.61	Koenigs–Knorr synthesis
+17.4	0.27	0.23	Liebeskind–Srogl cross-coupling
+16.7	0.14	0.12	Simmons–Smith reaction
+15.4	0.15	0.13	Sharpless–Jacobsen dihydroxylation ^w
+15.2	0.38	0.33	[2 + 2 + 2] Michael–Michael–1,6-Wittig ^x
+14.6	0.11	0.096	Pfitzner–Moffatt oxidation of primary alcohols to aldehydes ^y
+13.8	0.33	0.29	Reimer–Tiemann reaction
+13.8	0.33	0.29	Hoch–Campbell aziridine synthesis (R = Ph)
+12.5	0.63	0.56	aniline synthesis via potassium amide
+11.1	0.7	0.63	Dornow–Wiehler isoxazole synthesis
+11.1	0.3	0.27	Stevens rearrangement (bromide salt)
+10.5	0.63	0.57	Malaprade oxidation
+10.0	0.88	0.8	Willgerodt reaction
+9.8	0.56	0.51	Chichibabin pyridine synthesis
+9.4	0.35	0.32	Friedel–Crafts acylation (phenyl and Cl case, AlCl ₃ in stoichiometric amounts)
+7.1	0.75	0.7	Marshalk reaction
+5.6	0.38	0.36	Lemieux–Johnson oxidation of olefins to 1,2-diols ^z
+4.5	0.23	0.22	synthesis of acetates via oxymercuration ^{aa}
-7.2	0.64	0.69	hydroboration–borane rearrangement
-8.3	0.22	0.24	Bamford–Stevens oxidation of hydrazones
-9.1	0.3	0.33	Kulinkovich reaction (with amides)
-14.0	0.37	0.43	Harries ozonolysis of olefins with reductive workup ^{bb}
-16.0	0.42	0.5	lithium aluminum hydride reduction of ketones
-19.0	0.17	0.21	Oppenauer oxidation ^{cc}
-19.8	0.077	0.096	Pfitzner–Moffatt oxidation of primary alcohols to aldehydes ^{dd}
-20.0	0.24	0.3	Lithium aluminum hydride reduction of methyl esters
-25.0	0.21	0.28	Grignard reaction (X = Br, ketone substrate)
-26.1	0.17	0.23	Boyle–Sims oxidation
-29.2	0.17	0.24	Tiemann rearrangement ^{ee}
-30.0	0.14	0.2	Rubottom oxidation
-30.4	0.39	0.56	Borohydride reduction
-32.6	0.31	0.46	Borch reduction (with water)
-34.0	0.031	0.047	Fleming oxidation
-36.4	0.14	0.22	Epoxidation of olefins with <i>m</i> -CPBA
-39.1	0.14	0.23	Baeyer–Villiger reaction with <i>m</i> -CPBA
-39.5	0.26	0.43	Harries ozonolysis of olefins with reductive workup ^{ff}
-42.9	0.12	0.21	Oppenauer oxidation ^{gg}
-46.8	0.25	0.47	Suzuki coupling (Z = I)
-50.0	0.14	0.28	Meerwein–Ponndorf–Verley reduction
-50.9	0.27	0.55	Suzuki coupling (Z = Br)

Table 11 (Continued)

% change in (AE) _{min}	(AE) _{min} with recycling byproducts	(AE) _{min} without recycling byproducts	named reaction
-54.5	0.3	0.66	Suzuki coupling (Z = Cl)
-60.5	0.17	0.43	Harries ozonolysis of olefins with reductive workup ^{hh}
-65.1	0.15	0.43	Harries ozonolysis of olefins with reductive workup ⁱⁱ
-65.2	0.073	0.21	Oppenauer oxidation ^{jj}
-69.8	0.13	0.43	Harries ozonolysis of olefins with reductive workup ^{kk}
-74.8	0.053	0.21	Oppenauer oxidation ^{ll}

^a Recycling reactions: $\text{OsO}_2(\text{OH})_2 + \text{H}_2\text{O}_2 \rightarrow \text{OsO}_4 + 2\text{H}_2\text{O}$ and $\text{OsO}_4 + \text{RNH}_2 \rightarrow \text{R}-\text{NOsO}_3 + \text{H}_2\text{O}$. ^b Recycling reaction: $4\text{HOAc} + \text{O}_2 + 2\text{Hg} \rightarrow 2\text{Hg}(\text{OAc})_2 + 2\text{H}_2\text{O}$ ($\text{PdCl}_2[\text{PPh}_3]_2$ catalyst). ^c Recycling reactions: $2\text{OsO}_2 + 4\text{H}_2\text{O}_2 \rightarrow 2\text{OsO}_4 + 4\text{H}_2\text{O}$ and $\text{NaIO}_3 + \text{H}_2\text{O}_2 \rightarrow \text{NaIO}_4 + \text{H}_2\text{O}$. ^d Recycling reaction: $1,4\text{-dihydroxybenzene} + \frac{1}{2}\text{O}_2 \rightarrow p\text{-benzoquinone} + \text{H}_2\text{O}$. ^e Recycling reaction: $2\text{OsO}_2 + 4\text{H}_2\text{O}_2 \rightarrow 2\text{OsO}_4 + 4\text{H}_2\text{O}$. ^f Recycling reactions: $\text{OsO}_2(\text{OH})_2 + \text{NMO} \rightarrow \text{OsO}_4 + \text{NMM}$ and $\text{OsO}_4 + \text{RNH}_2 \rightarrow \text{R}-\text{NOsO}_3 + \text{H}_2\text{O}$. ^g Recycling reactions: $2\text{OsO}_2 + 4\text{H}_2\text{O}_2 \rightarrow 2\text{OsO}_4 + 4\text{H}_2\text{O}$, $2\text{OsO}_4 + 4\text{H}_2\text{O} + 2\text{K}_2[\text{OsO}_2(\text{OH})_4] + \text{CH}_3\text{COOH}$, and $\text{K}_4[\text{Fe}(\text{CN})_6] + \frac{1}{2}\text{Cl}_2 \rightarrow \text{K}_3[\text{Fe}(\text{CN})_6] + \text{KCl}$. ^h Recycling reaction: $1,4\text{-dihydroxybenzene} + \frac{1}{2}\text{MnO}_2 \rightarrow p\text{-benzoquinone} + \frac{1}{2}\text{Mn} + \text{H}_2\text{O}$. ⁱ Recycling reaction: $\text{Me}_2\text{S}=\text{O} + \text{H}_2 \rightarrow \text{Me}_2\text{S} + \text{H}_2\text{O}$. ^j Recycling reaction: $(^6\text{Bu})_3\text{SnCl} + \text{NaH} \rightarrow (^6\text{Bu})_3\text{SnH} + \text{NaCl}$. ^k Recycling reaction: $(^6\text{Bu})_3\text{SnCl} + \text{NaBH}_4 \rightarrow (^6\text{Bu})_3\text{SnH} + \text{NaCl} + \text{BH}_3$ or $(^6\text{Bu})_3\text{SnCl} + \text{LiAlH}_4 \rightarrow (^6\text{Bu})_3\text{SnH} + \text{AlH}_3 + \text{LiCl}$. ^l Recycling reaction: $\text{Cy}-\text{NHCONH}-\text{Cy} + \text{COCl}_2 + 2\text{NH}_3 \rightarrow \text{Cy}-\text{N}=\text{C}=\text{N}-\text{Cy} + \text{CO}_2 + 2\text{NH}_4\text{Cl}$. ^m Recycling reaction: $\text{Cy}-\text{NHCONH}-\text{Cy} + \text{P}_2\text{O}_5 \rightarrow \text{Cy}-\text{N}=\text{C}=\text{N}-\text{Cy} + 2\text{HPO}_3$. ⁿ Recycling reaction: $\text{O}=\text{PPh}_3 + \text{HCl} + \frac{1}{4}\text{Ti}(\text{iOPr})_4 \rightarrow \text{PPh}_3 + \text{H}_2\text{O} + \text{Me}_2\text{C}=\text{O} + \frac{1}{4}\text{TiCl}_4$. ^o Recycling reaction: $\text{O}=\text{PPh}_3 + \text{COCl}_2 + \frac{2}{3}\text{Al} \rightarrow \text{PPh}_3 + \text{CO}_2 + \frac{2}{3}\text{AlCl}_3$. ^p Recycling reaction: $\text{PhSO}_2\text{OH} + \frac{1}{3}\text{POCl}_3 \rightarrow \text{PhSO}_2\text{Cl} + \text{H}_3\text{PO}_4$. ^q Recycling reactions: $\text{OsO}_2 + 2\text{NMO} \rightarrow 2\text{NMM} + \text{OsO}_4$ and $\text{NaIO}_3 + \text{NMO} \rightarrow \text{NaIO}_4 + \text{NMM}$. ^r Recycling reaction: $\text{O}=\text{PPh}_3 + \text{HCl} + \frac{1}{4}\text{Ti}(\text{iOPr})_4 \rightarrow \text{PPh}_3 + \text{H}_2\text{O} + \text{Me}_2\text{C}=\text{O} + \frac{1}{4}\text{TiCl}_4$. ^s Recycling reaction: $\text{O}=\text{PPh}_3 + \text{COCl}_2 + \frac{2}{3}\text{Al} \rightarrow \text{PPh}_3 + \text{CO}_2 + \frac{2}{3}\text{AlCl}_3$. ^v Recycling reaction: $1,4\text{-dihydroxybenzene} + t\text{BuOOH} \rightarrow p\text{-benzoquinone} + t\text{BuOH} + \text{H}_2\text{O}$. ^w See footnote ^g. ^x Recycling reaction: $\text{O}=\text{PPh}_3 + \text{H}_2\text{O} + \text{HSiCl}_3 \rightarrow \text{PPh}_3 + \text{SiO}_2 + 3\text{HCl}$. ^y Recycling reaction: $\text{POCl}_3 + 4\text{NaOH} + \text{Cy}-\text{NHCONH}-\text{Cy} \rightarrow \text{Cy}-\text{N}=\text{C}=\text{N}-\text{Cy} + 2\text{H}_2\text{O} + \text{NaH}_2\text{PO}_4 + 3\text{NaCl}$. ^z Recycling reaction: $\text{OsO}_2 + 2\text{NMO} \rightarrow 2\text{NMM} + \text{OsO}_4$. ^{aa} Recycling reactions: $\text{Hg} + 2\text{HNO}_3 + \text{H}_2\text{O}_2 \rightarrow \text{Hg}(\text{NO}_3)_2 + 2\text{H}_2\text{O}$, $\text{Hg}(\text{NO}_3)_2 + 2\text{NaNO}_3 + \text{H}_2\text{O}$, and $\text{HgO} + 2\text{HOAc} \rightarrow \text{Hg}(\text{OAc})_2 + \text{H}_2\text{O}$. ^{bb} Recycling reaction: $(\text{Me}_2\text{S}=\text{O}) + \text{Zn} + \text{H}_2\text{O} \rightarrow \text{Me}_2\text{S} + \text{Zn}(\text{OH})_2$. ^{cc} Recycling reaction: $1,4\text{-dihydroxybenzene} + \frac{1}{2}\text{H}_2\text{O}_2 + \frac{1}{2}\text{I}_2 \rightarrow p\text{-benzoquinone} + \text{H}_2\text{O} + \text{HI}$. ^{dd} Recycling reaction: $\text{Cy}-\text{NHCONH}-\text{Cy} + \text{Ts}_2\text{O} \rightarrow \text{Cy}-\text{N}=\text{C}=\text{N}-\text{Cy} + 2\text{TsOH}$. ^{ee} Recycling reaction: $\text{PhSO}_3\text{H} + (^6\text{Bu})_4\text{NCl} \rightarrow \text{PhSO}_2\text{Cl} + (^6\text{Bu})_4\text{NOH}$. ^{ff} Recycling reaction: $(\text{Me}_2\text{S}=\text{O}) + \frac{1}{2}\text{S}=\text{PBr}_3 \rightarrow \text{Me}_2\text{S} + \frac{1}{2}\text{PBr}_3 + \frac{1}{2}\text{SO}_2$. ^{gg} Recycling reaction: $1,4\text{-dihydroxybenzene} + \text{NaIO}_4 \rightarrow p\text{-benzoquinone} + \text{NaIO}_3$. ^{hh} Recycling reaction: $(\text{Me}_2\text{S}=\text{O}) + \text{Zn} + \text{Br}(\text{CH}_2)_3\text{Br} \rightarrow (\text{Me}_2\text{S} + \text{H}_2\text{O} + \text{ZnBr}_2 + \text{CH}_2=\text{CH}-\text{CH}=\text{CH}_2$. ⁱⁱ Recycling reaction: $(\text{Me}_2\text{S}=\text{O}) + 2\text{NaI} + \text{H}_2\text{O} \rightarrow \text{Me}_2\text{S} + \text{I}_2 + 2\text{NaOH}$. ^{jj} Recycling reaction: $1,4\text{-dihydroxybenzene} + (\text{AcO})_2\text{IPhCONHMe} \rightarrow p\text{-benzoquinone} + \text{IPhCONHMe} + 2\text{HOAc}$. ^{kk} Recycling reaction: $(\text{Me}_2\text{S}=\text{O}) + 2\text{TiCl}_3 + 2\text{HCl} \rightarrow \text{Me}_2\text{S} + 2\text{TiCl}_4 + \text{H}_2\text{O}$. ^{ll} Recycling reaction: $1,4\text{-dihydroxybenzene} + \text{Pb}(\text{OAc})_2 + (\text{AcO})_2\text{IPh} \rightarrow p\text{-benzoquinone} + \text{PhI} + \text{Pb} + 2\text{HOAc}$.

Table 12. Summary of overall atom economy gains and losses for combined named sequences and byproduct recycling reactions under the assumption of 100% yields for all reactions under stoichiometric conditions

% change in (AE) _{min}	(AE) _{min} with recycling byproducts	(AE) _{min} without recycling byproducts	named sequence
+87.2	0.88	0.47	Bischler–Napieralski reaction
+86.4	0.11	0.059	Mitsunobu reaction ^a
+85.2	0.5	0.27	Claisen–Ireland rearrangement
+66.7	0.12	0.072	Wittig reaction ^b
+53.3	0.23	0.15	Wohl degradation (C6 sugars)
+51.7	0.44	0.29	Nenitzescu–Praill synthesis
+50.0	0.21	0.14	Wohl degradation (C5 sugars)
+42.9	0.2	0.14	Pfau–Plattner azulene synthesis
+41.4	0.41	0.29	Haworth phenanthrene synthesis
+41.2	0.24	0.17	Zemplen degradation (C5 sugars)
+38.9	0.1	0.072	Wittig reaction ^c
+38.9	0.25	0.18	Zemplen degradation (C6 sugars)
+36.4	0.15	0.11	Barbier–Wieland reaction (target: carboxylic acid)
+34.3	0.09	0.067	Schlosser modification of Wittig reaction ^d
+31.3	0.088	0.067	Schlosser modification of Wittig reaction ^e
+22.4	0.082	0.067	Schlosser modification of Wittig reaction ^f
+20.0	0.18	0.15	Barbier–Wieland reaction (target: ketone)
+15.9	0.51	0.44	Traube synthesis
+9.8	0.56	0.51	Chichibabin pyridine synthesis
+8.8	0.099	0.091	Doering–Laflamme synthesis
+6.0	0.53	0.5	Kolbe synthesis of aspirin
+3.7	0.056	0.054	Murahashi reaction ^g
0.0	0.059	0.059	Mitsunobu reaction ^h
-1.9	0.053	0.054	Murahashi reaction ⁱ
-7.1	0.13	0.14	Meyers aldehyde synthesis (X = Br)

^a Recycling reaction: $\text{Cl}_2 + \text{EtOOOC}-\text{NHNH}-\text{COOEt} \rightarrow \text{EtOOOC}-\text{N}=\text{N}-\text{COOEt} + 2\text{HCl}$ or $\text{HNO}_3 + \text{EtOOOC}-\text{NHNH}-\text{COOEt} \rightarrow \text{EtOOOC}-\text{N}=\text{N}-\text{COOEt} + \text{HNO}_2 + \text{H}_2\text{O}$. ^b Recycling reaction: $\text{O}=\text{PPh}_3 + \text{COCl}_2 + \frac{2}{3}\text{Al} \rightarrow \text{PPh}_3 + \text{CO}_2 + \frac{2}{3}\text{AlCl}_3$ or $\text{O}=\text{PPh}_3 + \text{HCl} + \frac{1}{4}\text{Ti}(\text{iOPr})_4 \rightarrow \text{PPh}_3 + \text{H}_2\text{O} + \text{Me}_2\text{C}=\text{O} + \frac{1}{4}\text{TiCl}_4$. ^c Recycling reaction: $\text{O}=\text{PPh}_3 + \text{H}_2\text{O} + \text{HSiCl}_3 \rightarrow \text{PPh}_3 + \text{SiO}_2 + 3\text{HCl}$. ^d Recycling reaction: $\text{O}=\text{PPh}_3 + \text{COCl}_2 + \frac{2}{3}\text{Al} \rightarrow \text{PPh}_3 + \text{CO}_2 + \frac{2}{3}\text{AlCl}_3$. ^e Recycling reaction: $\text{O}=\text{PPh}_3 + \text{COCl}_2 + \frac{2}{3}\text{Al} \rightarrow \text{PPh}_3 + \text{CO}_2 + \frac{2}{3}\text{AlCl}_3$. ^f Recycling reaction: $\text{O}=\text{PPh}_3 + \text{H}_2\text{O} + \text{HSiCl}_3 \rightarrow \text{PPh}_3 + \text{SiO}_2 + 3\text{HCl}$. ^g Recycling reaction: $\text{O}=\text{PPh}_3 + \text{COCl}_2 + \frac{2}{3}\text{Al} \rightarrow \text{PPh}_3 + \text{CO}_2 + \frac{2}{3}\text{AlCl}_3$ or $\text{O}=\text{PPh}_3 + \text{HCl} + \frac{1}{4}\text{Ti}(\text{iOPr})_4 \rightarrow \text{PPh}_3 + \text{H}_2\text{O} + \text{Me}_2\text{C}=\text{O} + \frac{1}{4}\text{TiCl}_4$. ^h Recycling reaction: $\text{PhI}(\text{OAc})_2 + \text{EtOOOC}-\text{NHNH}-\text{COOEt} \rightarrow \text{EtOOOC}-\text{N}=\text{N}-\text{COOEt} + \text{PhI} + 2\text{HOAc}$. ⁱ Recycling reaction: $\text{O}=\text{PPh}_3 + \text{H}_2\text{O} + \text{HSiCl}_3 \rightarrow \text{PPh}_3 + \text{SiO}_2 + 3\text{HCl}$.

The corresponding E_{\max} values are 12.93 and 7.64 for $x = 3$ (all R groups are H). These immediately translate to $(\text{AE})_{\min}$ values of 0.072 and 0.12, respectively. Hence, recycling

triphenylphosphine oxide increases $(\text{AE})_{\min}$ for the Wittig reaction by about 67%. From the analysis of recycling byproducts given previously,¹ the general expression for the

net change in $(AE)_{min}$ under the assumption of 100% yields for all reactions under stoichiometric conditions is given by

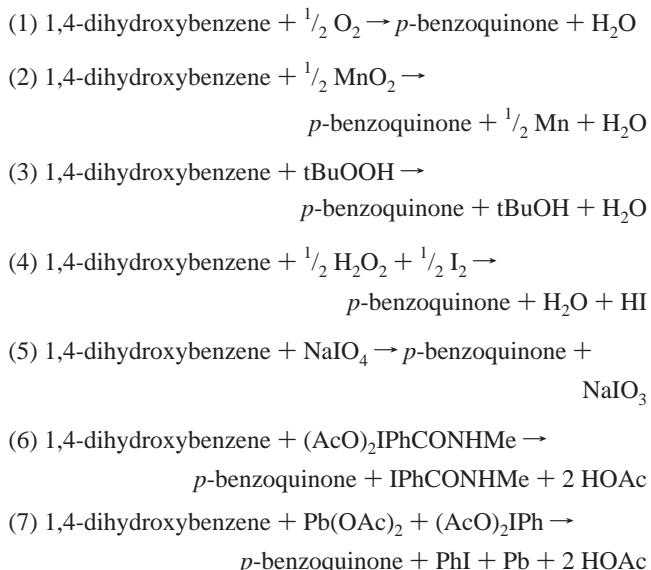
$$\Delta(\text{AE})_{\min} = \frac{[(\text{AE})_{\min}]^2(a-c)}{p - (\text{AE})_{\min}(a-c)} \quad (28)$$

where a , c , and p are the molecular weights of recovered reagent, the reagent(s) used in the recycling reaction, and the target product. The net change in $(AE)_{\min}$ is positive only if $a > c$ and $p > (AE)_{\min}$ ($a - c$); otherwise it is negative. If the yields of the forward, ϵ , and recycling reactions, ϵ^* , are accounted for, then eq 28 is modified to

$$\Delta(\text{RME})_{\min} = \frac{\epsilon^2 [(\text{AE})_{\min}]^2 (a\epsilon^* - c)}{p - (\text{AE})_{\min} \epsilon (a\epsilon^* - c)} = \frac{[(\text{RME})_{\min}]^2 (a\epsilon^* - c)}{p - (\text{RME})_{\min} (a\epsilon^* - c)} \quad (29)$$

where the change is now with respect to the reaction mass efficiency. Again, a positive gain in $(RME)_{\min}$ is possible only if $a\epsilon^* > c$ and $p > (RME)_{\min}(a\epsilon^* - c)$; otherwise it is negative.

These kinds of calculations can be used to quickly screen out recycling reaction options that would result in clear negative changes in $(AE)_{min}$. For example, the Oppenauer oxidation waste byproduct 1,4-dihydroxybenzene may be reoxidized back to *p*-benzoquinone by a variety of oxidizing agents according to the following published methods (for references, see Table R1, Supporting Information):



as shown in footnotes *d*, *h*, *v*, *cc*, *gg*, *jj*, and *ll* in Table 11. The present analysis shows that recycling options 4–7 can be eliminated since they yield substantial negative changes in $(AE)_{min}$ (−19%, −43%, −65%, and −75%, respectively) when combined with the original Oppenauer oxidation reaction; whereas, options 1–3 yield overall minimum atom economy gains of +200%, +91%, and +19%, respectively. In cases where all available recycling options give negative changes in $(AE)_{min}$ then one is faced with the challenge of inventing a possible recycling option that meets the criterion

that the overall change in $(AE)_{min}$ is positive if the disposal option is to be avoided. The merit of the present analysis is that it can direct a chemist's decision to choose among several options with some degree of confidence.

7. Application of Database to Diversity-Oriented Synthesis and Multicomponent Reaction Discovery

We now demonstrate how the constructed database of named organic reactions may be used in a diversity-oriented approach¹⁵ to discover new multicomponent reactions. Table 13 summarizes the number of occurrences and frequencies of various Markush fragments appearing in skeletal building reactions, namely, multicomponent reactions, carbon–carbon bond-forming reactions, non-carbon–carbon bond-forming reactions, and condensations. One can see that the top 10 most abundant fragments are aldehydes, amines, alkynes, alkenes, alkyl halides, ketones, α -H ketones, isocyanides, nitriles, and α,β -unsaturated ketones or esters. This list is not surprising, given that these are the usual building blocks used to develop combinatorial libraries.¹⁶ Table 14 summarizes results of simple combinatorics to determine the number of possible combinations of three- and four-component reactions as a function of the size of the Markush fragment library.

Although the combinatorial analysis shown in Table 14 indicates a combinatorial explosion of possibilities, it does not necessarily follow that all of the combinations lead to productive multicomponent reactions. There may be combinations that lead to too many possible reaction products which would make practical control of the intended target product difficult to implement, or worse still, the combination may not have the required number of electrophile–nucleophile pairings of active sites on the components to yield a product that contains all fragments. For any one of the entries listed in Table 14 it is difficult to predict the fraction of combinations that are, in fact, viable multicomponent reactions that could be tested in the laboratory. At present, there is no method available which can rapidly screen out nonproductive combinations. It is possible to enumerate the vast array of combinations by computer methods, but it is a laborious process to sift through each of the possibilities to determine the chemical viability of the combinations. Although the “rules” of chemical reactivity based on the identification of electrophilic and nucleophilic sites on molecular structures are difficult to encode in machine language, the chemical intuition of a chemist still remains the best screening tool available at present. The task is analogous to constructing three- or four-letter words from an alphabet of many letters with the added difficulty that we do not have a lexicon or dictionary to check the sense of the words. In addition, there is also the further complexity to consider that the same combination of building blocks may yield different kinds of multicomponent products,

(15) (a) Burke, M. D.; Schreiber, S. L. *Angew. Chem., Int. Ed.* **2004**, *43*, 46. (b) Burke, M. D.; Berger, E. M.; Schreiber, S. L. *Science* **2003**, *302*, 613. (c) Schreiber, S. L. *Science* **2002**, *287*, 1094. (d) Burke, M. D.; Lalic, G. *Chem.*

(16) Armstrong, R. W.; Combs, A. P.; Tempest, P. A.; Brown, S. D.; Keating, T. A. *Acc. Chem. Res.* **1996**, 29, 123.

Table 13. Summary of occurrences and frequencies of Markush groups in reaction database

Structural Group	Structure	Number of Occurrences	Frequency (%)	Structural Group	Structure	Number of Occurrences	Frequency (%)
		in Database				in Database	
aldehyde		55	13.41	1,2-diketone		4	0.98
amine		34	8.29	barbiturate		3	0.73
alkyne		25	6.10	diazomethane		3	0.73
alkene		25	6.10	aryldiazonium salt		3	0.73
alkyl halide		18	4.39	Grignard or organocuprate reagent		3	0.73
ketone		18	4.39	organolithium reagent		3	0.73
α -H ketone		13	3.17	silylenol ether		3	0.73
isocyanide		11	2.68	urea		3	0.73
nitrile		10	2.44	indole		3	0.73
α,β -unsaturated ketone or ester		10	2.44	Meldrum's acid		3	0.73
alcohol		10	2.44	carbon dioxide		3	0.73
β -ketoester		9	2.20	allyltrimethylsilane		2	0.49
aryl		9	2.20	α -cyanoester		2	0.49
carbon monoxide		8	1.95	amide		2	0.49
1,3-diene		6	1.46	anhydride		2	0.49
acid chloride		6	1.46	cyclobutene		2	0.49
acid		5	1.22	isocyanate		2	0.49
allene		5	1.22	malonate		2	0.49
α -halo ketone		5	1.22	substituted pyridine		2	0.49
α -keto ester		5	1.22	arylboronic acid		2	0.49
aniline		5	1.22				
imine		5	1.22				
phenol		5	1.22				
aryl halide		4	0.98				
aryl iodide		4	0.98				
organoboronic acid		4	0.98				
vinyl ether		4	0.98				

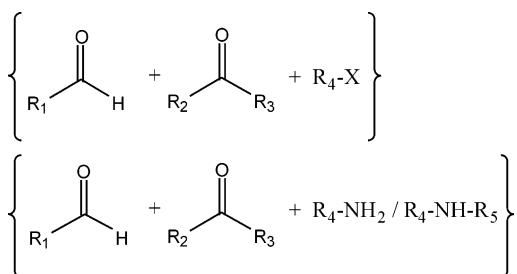
^a Frequency calculations based on a database of 410 reactions.

Table 14. Combinatorial analysis of possible three-component and four-component MCRs as a function of Markush fragment library sizes

library size of Markush fragments, <i>n</i>	number of possible three-component MCRs, <i>C</i> (<i>n</i> ,3) ^a	number of possible four-component MCRs, <i>C</i> (<i>n</i> ,4) ^b
7	35	35
8	56	70
9	84	126
10	120	210
15	455	1,365
20	1,140	4,845
25	2,300	12,650
30	4,060	27,405
35	6,545	52,360
40	9,880	91,390
45	14,190	148,995
50	19,600	230,300
100	161,700	3,921,225

^a Calculated using $C(n,3) = [n!/3!(n-3)!]$. ^b Calculated using $C(n,4) = [n!/4!(n-4)!]$.

Scheme 6



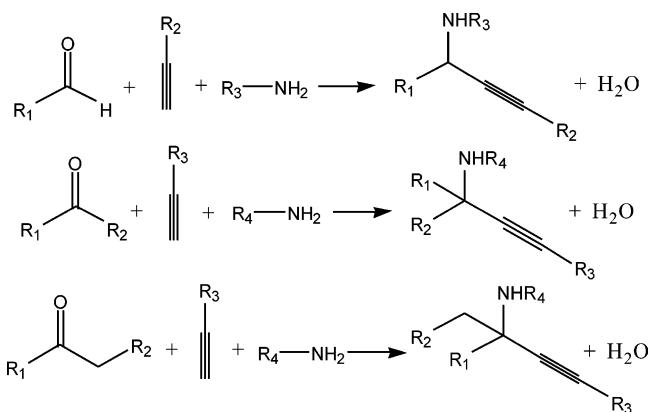
depending on the reaction conditions and/or the order of addition of substrates. Using the letter-word analogy this problem is equivalent to constructing more than one possible three- or four-letter word from the same set of three or four letters. As an example, we may form the words “eat”, “tea”, and “ate” from the three letters A, E, and T.

Nevertheless, despite these complexities, it is possible to come up with surprising results even with a small-size database of Markush fragments that can be enumerated and analyzed by hand. In this regard, a table of predicted three-component MCRs based on the top seven most frequently appearing Markush fragments in the database is given in the Supporting Information. Estimates of $(AE)_{min}$ and E_{max} are also included as well as conditions where order of addition of reagents matters with respect to the product outcome. Of the 35 total possible combinations, two of them, or about 6%, do not lead to viable multicomponent products for reasons mentioned above. These nonproductive combinations are shown in Scheme 6. Of the remaining 33 productive combinations, 21 have not been documented and would therefore be considered new, and 12 have recent literature precedence. These represent 60% and 34%, respectively, of the total number of 35 possible combinations. It is clear from these figures that, as other authors have pointed out,⁹ the domain of multicomponent reactions has only begun to be explored.

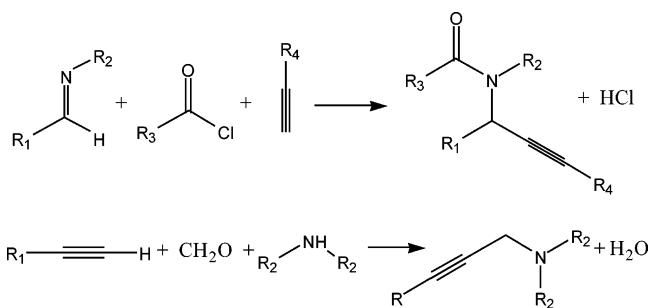
We now present the 12 predicted MCRs from this analysis that have literature precedence in Schemes 7–15.

The synthesis of propargylamines using metal-catalyzed multicomponent strategies as predicted by the above analysis

Scheme 7



Scheme 8



(see Scheme 7) have recently been reported¹⁷ including variants using an imine, acid chloride, and acetylene combination¹⁸ and an amine, formaldehyde, and acetylene combination¹⁹ as shown in Scheme 8.

The predicted MCR shown in Scheme 9 is reminiscent of the Stork enamine reaction.²⁰

The predicted MCRs shown in Scheme 10 are Mannich-type condensations when compared with the original Mannich reaction shown in the third entry.²¹

The predicted MCRs shown in Scheme 11 resemble the reported synthesis of pyrano- and furanoquinolines via tandem imination-[2 + 2] cycloaddition as given in the third entry.²²

The predicted MCR given in Scheme 12 is an aldol synthesis when R_4 is H.

The predicted MCRs shown in Scheme 13 are tandem aldol–Michael reactions. In the third and fourth entries there are two possible modes of reaction with respect to the olefin on the aldol adduct, depending on the electronic nature of the olefin substituent and whether the aldol product eliminates water. If the reaction is controlled to produce the aldol

(17) (a) Wei, C.; Li, C. *J. Am. Chem. Soc.* **2003**, *125*, 9584. (b) Gommerman, N.; Koradin, C.; Polborn, K.; Knochel, P. *Angew. Chem., Int. Ed.* **2003**, *42*, 5763. (c) Wei, C.; Li, Z.; Li, C. *J. Org. Lett.* **2003**, *5*, 4473. (d) Li, Z.; Wei, C.; Chen, L.; Varma, R. S.; Li, C. *J. Tetrahedron Lett.* **2004**, *45*, 2443. (e) Shi, L.; Tu, Y. Q.; Wang, M.; Zhang, F. M.; Fan, C. A. *Org. Lett.* **2004**, *6*, 1001. (f) Yadav, J. S.; Reddy, B. V. S.; Naveenkumar, V.; Rao, R. S.; Nagaiah, K. *New J. Chem.* **2004**, *28*, 335.

(18) Black, D. A.; Arndsten, B. A. *Org. Lett.* **2004**, *6*, 1107.

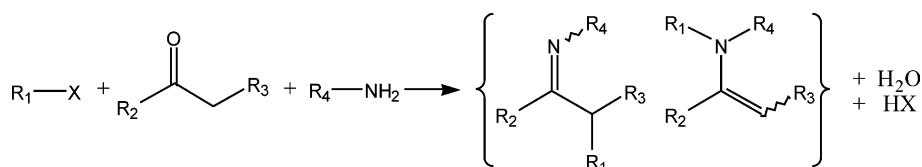
(19) Sharifi, A.; Farhangian, H.; Mohsenzadeh, F.; Naimi-Jamal, M. R. *Monatsh. Chem.* **2002**, *133*, 199.

(20) Stork, G.; Terrell, R.; Szmuszkovicz, J. *J. Am. Chem. Soc.* **1954**, *76*, 2029.

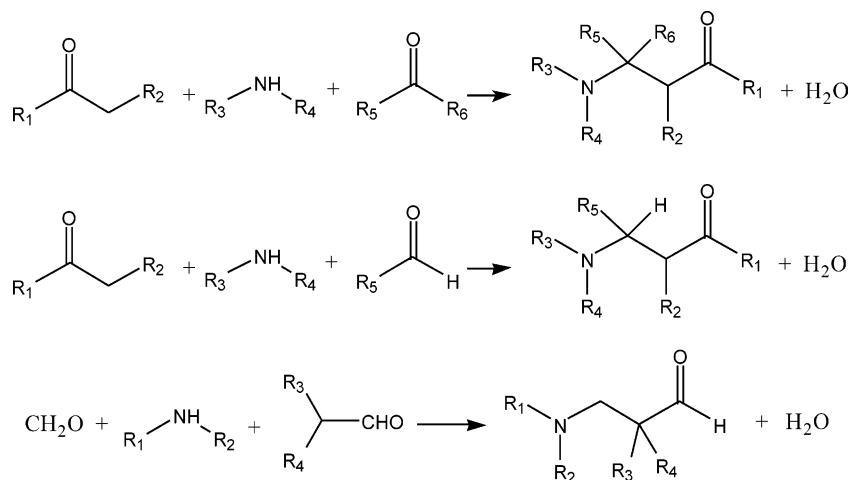
(21) Mannich, C.; Kroesche, I. *Arch. Pharm. (Weinheim, Ger.)* **1912**, *250*, 647.

(22) (a) Yadav, J. S.; Reddy, B. V. S.; Reddy, J. S. S.; Rao, R. S. *Tetrahedron Lett.* **2003**, *59*, 1599. (b) Ravindranath, N.; Ramesh, C.; Reddy, M. R.; Das, B. *Chem. Lett.* **2003**, *32*, 222.

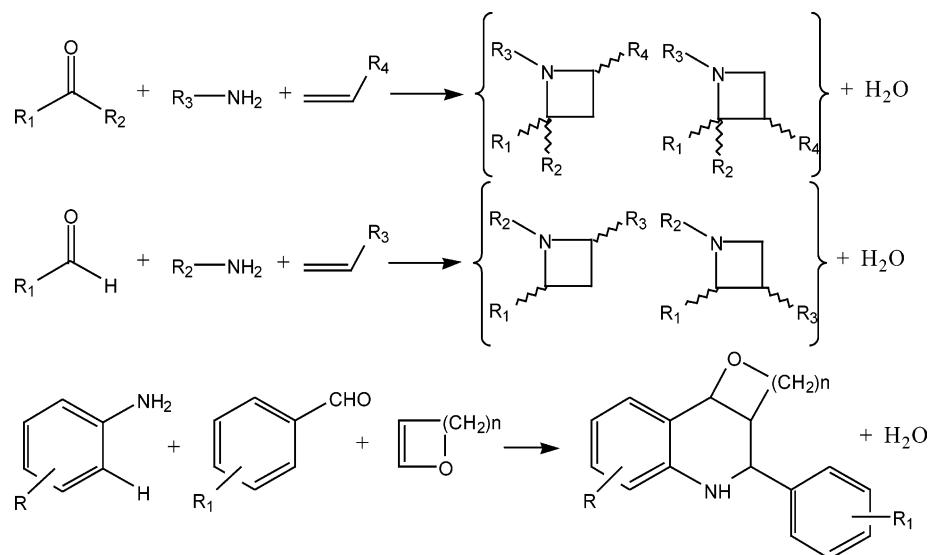
Scheme 9



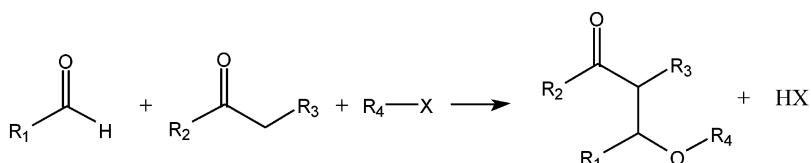
Scheme 10



Scheme 11



Scheme 12

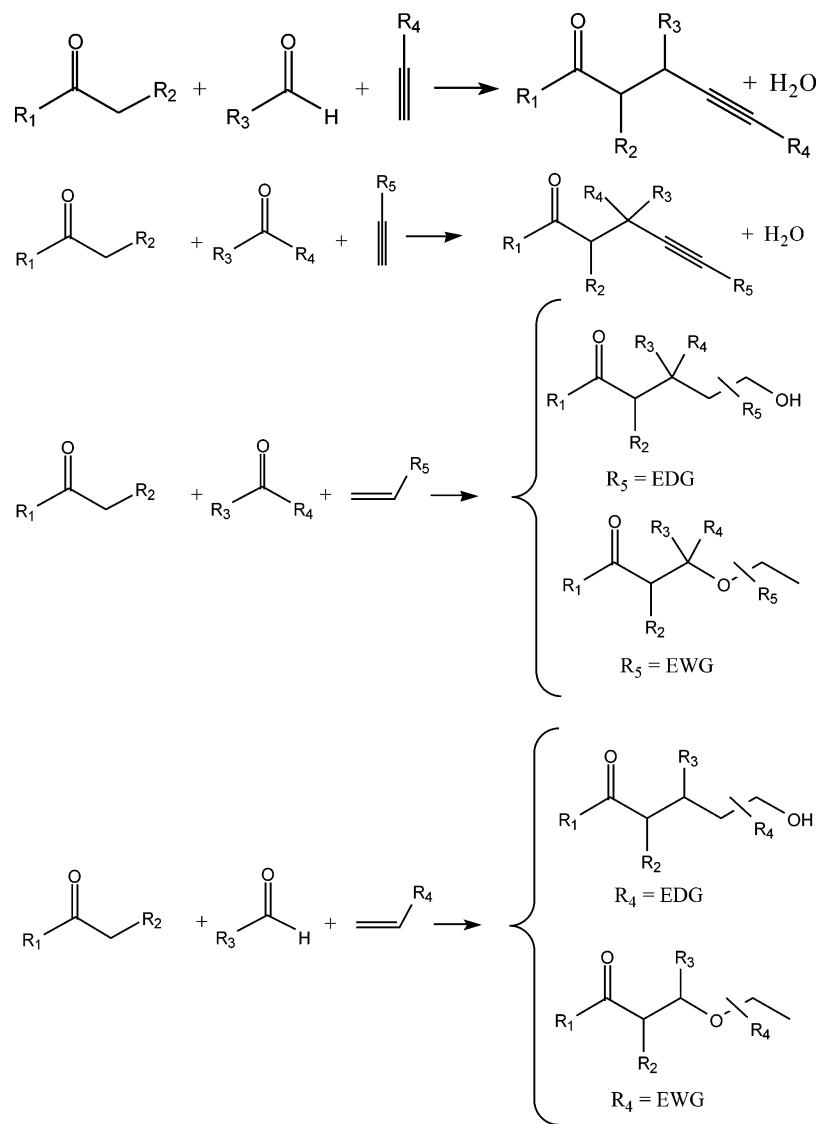


adduct without eliminating water, then it could act as a nucleophile toward an electrophilic alkene. On the other hand, if the originally formed aldol adduct eliminates water to yield the α,β -unsaturated ketone, then it can be attacked in a Michael fashion by the π -bond of a nucleophilic alkene. The reverse sequence, namely, a tandem Michael–aldol reaction was recently reported using the combination of an alkynyl ketone, ketone, and sulfide shown in Scheme 14.²³

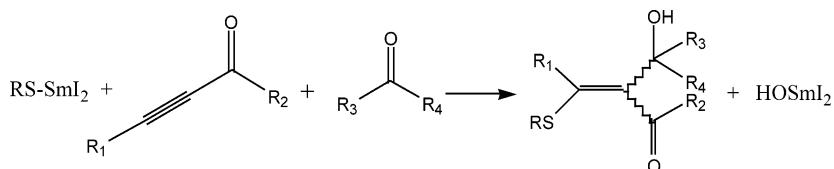
It is interesting to note that the same combination of fragments shown in entries 3 and 4 of Scheme 13 can also lead to tandem aldol–[2 + 2] products under conditions where the initially formed α,β -unsaturated ketone undergoes subsequent [2 + 2] cycloaddition with the olefin component as shown in Scheme 15.

(23) Zheng, X.; Yu, X.; Zhang, Y. *Synlett* 2003, 2062.

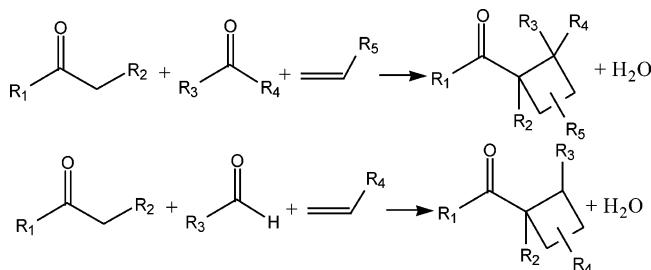
Scheme 13



Scheme 14



Scheme 15

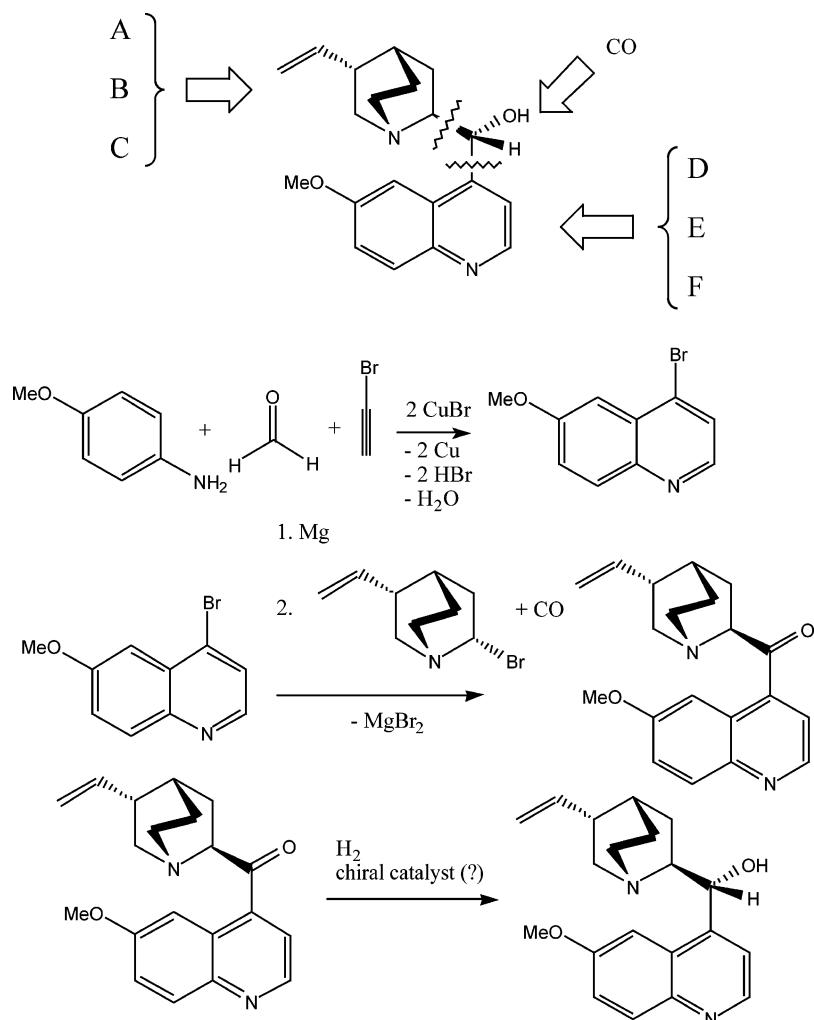


To make the diversity-oriented approach by the proposed method purposeful and useful in developing synthetic strategies it must be linked to a database of known complex structures. If substructures of complex molecules may be

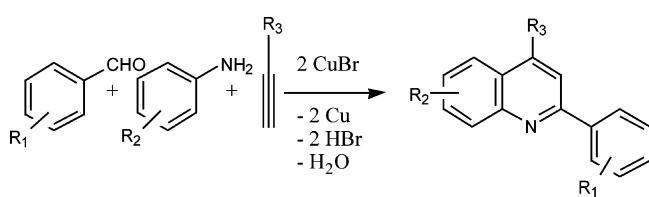
identified as entities that are, in fact, the products of multicomponent reactions predicted by the above database of Markush fragments, then this gives reason to pursue synthetic routes by the multicomponent approach. This mapping of databases will allow work done by the diversity-combinatorial approach to complement that of identifying complex targets existing in Nature.

We now demonstrate this idea by proposing an alternate total synthesis of *(-)*-quinine as shown in Scheme 16 that makes use of a *nested* multicomponent strategy. A nested multicomponent strategy represents the maximum degree of convergence possible. The first step to prepare the quinoline moiety by a three-component coupling is a known reaction as shown by the entry in Scheme 17.²⁴ The third step is a

Scheme 16



Scheme 17



redox reaction as in the final stereoselective oxidation step employed in the recent Stork synthesis of this target²⁵ except that here a stereoselective reduction is proposed which may or may not require a chiral catalyst. The second reaction is a new tandem Grignard–carbonylation–substitution sequence. The predicted overall atom economy and kernel environmental impact factor for this nested MCR strategy are 39.8% and 1.52, respectively. These values may be compared with an overall AE of about 8% and an overall E_{mw} value of about 11 for both linear Woodward–Rabe and Stork strategies.¹ It should be noted that both the Woodward–Rabe and Stork syntheses focused on the construction of the quinuclidine moiety, albeit by very different means, and that the quinoline

moiety was introduced as a single entity toward the end of the respective sequences. What is missing in the present proposed synthetic strategy is a three-component coupling reaction that can afford the quinuclidine moiety in a single step or tandem sequence. Indeed, it is this part of the structure which in fact presents the greatest challenge in the total synthesis of quinine. If such an MCR can be found, then the above estimates for the overall AE and E_{mw} values for this sequence would be upper and lower limits respectively, since it is likely that an MCR leading to the quinuclidine structure would generate its own waste byproducts. Given that the best-case scenario for the overall AE is about 5 times better than for the published procedures, it seems reasonable to suggest that such an atom economical strategy is worth pursuing.

8. Conclusions and Synthetic Challenges Paradigm

The quantitative results presented in this work clearly indicate that the “green” standard constraint is at first sight difficult to achieve but is not insurmountable as indicated by the analysis of the library of named organic reactions. From Figure 5 about 55% of all reactions considered in the present database would fall in the “green” zone if their minimum atom economies of 61.8% are coupled with high yields and high solvent recovery and are run under stoichio-

(24) Zhang, J.; Yang, W.; Song, L.; Cai, X.; Zhu, S. *Tetrahedron Lett.* **2004**, *45*, 5771.

(25) 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.

metric conditions so that their overall RME values are above 61.8%. However, particular categories of reactions are better performers than others as shown by the results presented in Figure 6. The overall aim, therefore, is to invent new chemical reactions that fall in the “green” zone by having intrinsic atom economies and reaction yields as high as possible, that have the greatest possibility of solvent recovery and/or elimination, and that are run without excess reagents. The suggestion put forward here is that the “greenness” criterion can in fact be rigorously quantified and is a bounded problem. Maintaining AE values above 0.618 is critical in increasing the chances that RME is also above 0.618; otherwise dramatic attenuations in RME result. It is important to note that the high atom economy and reaction mass efficiency metrics for the syntheses of target products have the highest priorities in the list of 12 Principles of Green Chemistry as advanced by Paul Anastas and John Warner.¹³ The rapid progress of new multicomponent reactions reported in the literature in the last 3 years is a testament that efforts are already underway to meet this challenge. Thus far such reports demonstrate success for isolated reactions. What is currently lacking is the integration of such multicomponent reactions as core strategies and themes in total synthesis design as put forward in the above illustrative example of the classic molecule quinine.

On the basis of the findings presented in this work and the previous report¹ one may summarize the quest for developing ideal efficient^{13,26,27} reactions and synthetic routes to important targets according to the following modern paradigm:

- (1) minimize the overall number of steps, preferably using nested MCRs as far as possible as the main strategy in a synthetic plan to a complex target;
- (2) maximize yield per step to be no lower than 75%;²⁸
- (3) maximize atom economy per step keeping $(AE)_{min}$ above the “golden” threshold of 61.8% and E_{max} below 0.618;
- (4) operate under stoichiometric conditions if possible, keeping SF equal to 1 for each reaction;
- (5) minimize overall E_{mw} factor and maximize overall AE in multistep syntheses: (a) by maximizing frequency of condensations, MCRs, rearrangements, carbon–carbon and non-carbon–carbon bond-forming reactions and (b) by minimizing frequency of substitutions (protecting group strategies) and redox reactions;
- (6) if forced to use oxidations, opt for hydrogen peroxide or molecular oxygen as oxidant;
- (7) if forced to use reductions, opt for molecular hydrogen as reductant;
- (8) devise electrochemical transformations;²⁹
- (9) devise catalytic methods where catalysts are recycled and reused;

(26) (a) Wender, P. A.; Handy, S. T.; Wright, D. L. *Chem. Ind. (London)* **1997**, 765. (b) Sheldon, R. A. *Chem. Ind. (London)* **1992**, 903.
 (27) (a) Matlack, A. *Green Chem.* **2003**, 5, G7. (b) Tundo, P.; Anastas, P.; Black, D. S.C.; Breen, J.; Collins, T.; Memoli, S.; Miyamoto, J.; Polyakoff, M.; Tumas, W. *Pure Appl. Chem.* **2001**, 72, 1207. (c) Clark, J. H. *Green Chem.* **1999**, 1, 1. (d) Datt, C. B.; Davis, M. E. *Ind. Eng. Chem.* **1994**, 33, 2887.
 (28) Hudlicky, T. *Chem. Rev.* **1996**, 96, 3.
 (29) (a) Matthews, M. A. *Pure Appl. Chem.* **2001**, 73, 1305. (b) Puetter, H.; Huber, G.; Merk, C. *GDCh-Monograph.* **2001**, 23, 81. (c) Yang, H.; Gu, Y.; Deng, Y.; Shi, F. *Chem. Commun.* **2002**, 274.

(10) devise regio- and stereoselective synthetic strategies particularly regio- and stereoselective MCRs;

(11) opt for solventless reactions, recycle solvents, or use benign solvents (e.g., ionic liquids, supercritical media);

(12) minimize energy demands such as heating, cooling, and carrying out reactions under pressure exceeding 1 atm; aim for reactions run under standard temperature and pressure conditions (room temperature and 1 atm).

The demonstration of this paradigm practically and economically to the synthesis of *existing* molecules of industrial importance that are currently made by traditional routes is probably the most effective way of convincing people engaged in the chemical enterprise to adopt it as a platform for the synthesis of molecules of future importance.

This report has introduced the concept of minimum atom economy and maximum environmental impact factor and applied it to a database of over 400 named organic reactions. Worst-case scenarios for the “green” performances of organic reactions in the database have been determined by using balanced chemical equations written in general form with Markush structures. A “golden” atom economical threshold has been proposed as the means of gauging the true “greenness” of chemical reactions, and various reaction classes have been analyzed and ranked according to this criterion. Associated probabilities for achieving this criterion for various reaction conditions have been evaluated. An assessment of the viability of recycling reaction byproducts based on available literature methods is presented with respect to net gains or losses in overall atom economy. The use of the database in a diversity–combinatorial approach for the discovery of new multicomponent reactions has been demonstrated. The “greenness” criterion for chemical reactions is rigorously defined and shown to be a bounded problem.

Acknowledgment

Maija Elina Lukkari, an undergraduate exchange student from the University of Helsinki, is thanked for pointing out the behaviour of the general form of the environmental impact factor when applied to reactions yielding multiple products containing variable functional groups. Prof. C. J. Li of McGill University is thanked for preliminary discussions regarding the use of the database to predict MCRs. Parts of this work were presented at the Canada-US 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.

Supporting Information Available

A database of more than 400 named organic reactions in the form of tables (Tables D1–D10) grouped according to multicomponent reactions, carbon–carbon bond-forming reactions, non-carbon–carbon bond-forming reactions, condensations, oxidations with respect to substrate, reductions

with respect to substrate, rearrangements, substitutions, fragmentations and eliminations, and sequences. Each reaction is generalized using Markush structures, chemically balanced accounting for all byproducts, and analyzed to determine the minimum atom economy and maximum environmental impact factor. Tables R1 and R2 list literature byproduct recycling reactions with yields of at least 90% applicable to named reactions where conversion of byprod-

ucts to starting reagents may be deemed viable. Also included in the database is a list of 33 predicted MCRs constructed from the seven most frequently occurring Markush fragments in the database. This information is available free of charge via the Internet at <http://www.pubs.acs.org>.

Received for review February 8, 2005.

OP050014V