Supplementary Information for

The Importance of Indirect Hotspots when Prioritizing Research in Green Chemical Synthesis

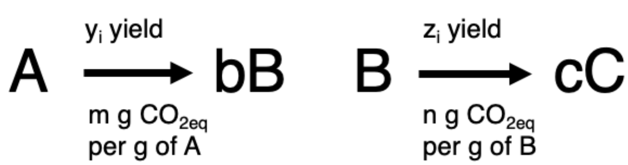
Philip G. Jessop and Alex R. MacDonald

Contents

S.1 Assumptions

Here we take a generic two-step reaction sequence (Scheme S1) and define the following variables:

starting mass of A = a (in grams)
stoichiometric coefficient of step 1 = b
stoichiometric coefficient of step 2 = c
molar masses of A, B, and C are M_A, M_B, and M_C
yield of step 1 (before modification) = y_i
yield of step 1 (after BY1 modification) = y_f
yield of step 2 (before modification) = z_i
yield of step 2 (after BY2 modification) = z_f
harm of step 1 per g A (before modification) = m (g CO_2 eq. per g A)
harm of step 1 per g A (after G1 modification) = xm (g CO_2 eq. per g A)
harm of step 2 per g B (before modification) = n (g CO_2 eq. per g B)
yield of step 2 per g A (after G2 modification) = xn (g CO_2 eq. per g B)



Scheme S1. A two-step reaction with variable amounts of harm and variable reaction stoichiometry.

We also make the following assumptions:

• The harm of any step includes the harm of making, using, and/or disposing of the reagents (except the starting material), byproducts, side products, solvents, and energy.

• The harm is shown here to be global warming, in units of g of CO2 equivalents, but could equally well be any other harm with the appropriate units.

• The harm of generating the first starting material "A" is not included in the analysis. "A" could be a chemical feedstock, a farmer's field, or even a mineral deposit.

• The harm of a step includes the harm of post-reaction separations/purifications. However, it is possible for the second step to be a separation/purification rather than a reaction. In such a situation $c = 1$.

S.2 Base case (no modifications)

<u>a) step 1</u>

starting mass of A = a (grams)

starting moles of A = a/M_A

yield of B = $y_i ab/M_A$ (moles)

yield of B = $y_i abM_B/M_A$ (grams)

harm of step 1 = am (g of CO_2 equivalents)

<u>b) step 2</u>

starting mass of B = $y_i abM_B/M_A$ (grams)

starting moles of B = $y_i ab/M_A$ (moles)

yield of C = $y_i z_i abc/M_A$ (moles)

yield of C = $y_i z_i abcM_C/M_A$ (grams)

harm of step 2 = $ny_i abM_B/M_A$ (g of CO_2 equivalents)

<u>c) total harm</u>

total harm = $am + ny_i abM_B/M_A$ (g of CO_2 equivalents)

total harm per g of C = $(am + ny_i abM_B/M_A)/(y_i z_i abcM_C/M_A)$

$= (m + ny_i bM_B/M_A)/(y_i z_i bcM_C/M_A)$

$= (mM_A + ny_i bM_B)/(y_i z_i bcM_C)$ (g of CO_2 equivalents per g of C)

<u>d) % harm per step</u>

% harm of step #1 = $am/(am + ny_i abM_B/M_A)$

$= m/(m + ny_i bM_B/M_A)$

$= mM_A/(mM_A + ny_i bM_B)$

% harm of step #2 = $(ny_i abM_B/M_A)/(am + ny_i abM_B/M_A)$

$= (ny_i abM_B)/(amM_A + ny_i abM_B)$

$= (ny_i bM_B)/(mM_A + ny_i bM_B)$

S.3 G1 Strategy

If we change the harm of the 1st step by a factor of x (where x is preferably less than 1, where an x of 1 means no change and an x of 0.1 means a 90% reduction in harm), then:

harm of step 1 = xam (g of CO_2 equivalents)
harm of step 2 = ny_iabM_B/M_A (g of CO_2 equivalents)
total harm = $xam + ny_iabM_B/M_A$ (g of CO_2 equivalents)
total harm per g of C = $(xam + ny_iabM_B/M_A)/(y_iz_iabcM_C/M_A)$
$= (xm + ny_ibM_B/M_A)/(y_iz_ibcM_C/M_A)$
$= (xmM_A + ny_ibM_B)/(y_iz_ibcM_C)$ (g of CO_2 equivalents per g of C)

$$\text{\% reduction in total harm per g C} = \frac{\text{total harm per g C (before)} - \text{total harm per g C (after)}}{\text{total harm per g C (before)}}$$

$$= \frac{(mM_A + ny_ibM_B)/(y_iz_ibcM_C) - (xmM_A + ny_ibM_B)/(y_iz_ibcM_C)}{(mM_A + ny_ibM_B)/(y_iz_ibcM_C)}$$

$$= \frac{(mM_A + ny_ibM_B) - (xmM_A + ny_ibM_B)}{mM_A + ny_ibM_B}$$

$$= \frac{(1-x)mM_A}{mM_A + ny_ibM_B}$$

S.4 BY1 Strategy

If we improve the yield of the 1st step to y_f, then:
starting mass of B = y_fabM_B/M_A (grams)
starting moles of B = y_fab/M_A (moles)
yield of C = y_fz_iabc/M_A (moles)
yield of C = $y_fz_iabcM_C/M_A$ (mass)
harm of step 2 = ny_fabM_B/M_A (g of CO_2 equivalents)
total harm = $am + ny_fabM_B/M_A$ (g of CO_2 equivalents)
total harm per g of C = $(am + ny_fabM_B/M_A) / (y_fz_iabcM_C/M_A)$
$= (m + ny_fbM_B/M_A) / (y_fz_ibcM_C/M_A)$
$= (mM_A + ny_fbM_B) / (y_fz_ibcM_C)$ (g of CO_2 equivalents per g of C)

$$\text{\% reduction in total harm per g C} = \frac{\text{total harm per g C (before)} - \text{total harm per g C (after)}}{\text{total harm per g C (before)}}$$

$$= \frac{(mM_A + ny_ibM_B)/(y_iz_ibcM_C) - (mM_A + ny_fbM_B)/(y_fz_ibcM_C)}{(mM_A + ny_ibM_B)/(y_iz_ibcM_C)}$$

$$= \frac{(mM_A + ny_ibM_B)/y_i - (mM_A + ny_fbM_B)/y_f}{(mM_A + ny_ibM_B)/y_i}$$

$$= \frac{(mM_A + ny_ibM_B) - (mM_A + ny_fbM_B)(y_i/y_f)}{mM_A + ny_ibM_B}$$

$$= 1 - \frac{y_i(mM_A + ny_fbM_B)}{y_f(mM_A + ny_ibM_B)}$$

$$= 1 - \frac{y_imM_A + ny_iy_fbM_B}{y_fmM_A + ny_iy_fbM_B}$$

S.5 G2 Strategy

If we change the harm of the 2nd step by a factor of x (where x is preferably less than 1, where an x of 1 means no change and an x of 0.1 means a 90% reduction in harm), then:

harm of step 1 = am (g of CO_2 equivalents)
harm of step 2 = xny_iabM_B/M_A (g of CO_2 equivalents)
total harm = $am + xny_iabM_B/M_A$ (g of CO_2 equivalents)
total harm per g of C = $(am + xny_iabM_B/M_A)/(y_iz_iabcM_C/M_A)$
$= (m + xny_ibM_B/M_A)/(y_iz_ibcM_C/M_A)$
$= (mM_A + xny_ibM_B)/(y_iz_ibcM_C)$ (g of CO_2 equivalents per g of C)

$$\text{\% reduction in total harm per g C} = \frac{\text{total harm per g C (before) - total harm per g C (after)}}{\text{total harm per g C (before)}}$$

$$= \frac{(mM_A + ny_ibM_B)/(y_iz_ibcM_C) - (mM_A + xny_ibM_B)/(y_iz_ibcM_C)}{(mM_A + ny_ibM_B)/(y_iz_ibcM_C)}$$

$$= \frac{(mM_A + ny_ibM_B) - (mM_A + xny_ibM_B)}{mM_A + ny_ibM_B}$$

$$= \frac{(1-x)ny_ibM_B}{mM_A + ny_ibM_B}$$

S.6 BY2 Strategy

If we improve the yield of the 2nd step to z_f, then:
yield of C = $y_iz_fabcM_C/M_A$ (mass)
total harm = $am + ny_iabM_B/M_A$ (g of CO_2 equivalents)
total harm per g of C = $(am + ny_iabM_B/M_A) / (y_iz_fabcM_C/M_A)$
$= (m + ny_ibM_B/M_A) / (y_iz_fbcM_C/M_A)$
$= (mM_A + ny_ibM_B) / (y_iz_fbcM_C)$ (g of CO_2 equivalents per g of C)

$$\% \text{ reduction in total harm per g C} = \frac{\text{total harm per g C (before)} - \text{total harm per g C (after)}}{\text{total harm per g C (before)}}$$

$$= \frac{(mM_A + ny_ibM_B)/(y_iz_ibcM_C) - (mM_A + ny_ibM_B)/(y_iz_fbcM_C)}{(mM_A + ny_ibM_B)/(y_iz_ibcM_C)}$$

$$= \frac{(mM_A + ny_ibM_B)/z_i - (mM_A + ny_ibM_B)/z_f}{(mM_A + ny_ibM_B)/z_i}$$

$$= \frac{(mM_A + ny_ibM_B) - (mM_A + ny_ibM_B)(z_i/z_f)}{mM_A + ny_ibM_B}$$

$$= 1 - \frac{z_i}{z_f}$$