Supporting Information For

Mathematical Relationships of Individual Stereocenter *er* Values to *dr* Values

Brenden Herrera[‡], Chung-Yon Lin[§], Alexandra Wright[⊥], Sarah Moor[‡] and Eric V. Anslyn[*,‡]

[‡]Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States
[§]Department of Molecular Medicine, Scripps Research Institute, La Jolla, CA 92037, United States
[⊥]Office of the Registrar, The University of Texas at Austin, Austin, Texas 78712, United States

Table of Contents

I. Derivation of a mathematical relationship relating enantiomeric ratios (*er*) at individual stereocenters to a global diastereomeric ratio (*dr*).

We start by defining the enantiomeric ratio (*er*) as *R* relative to *S* (Eq.S1), the alcohol (er_2) and amine (er_3) individual stereocenters in the manner that current optical methods would measure them (Eqs.S2 and S3, respectively), and the *dr* as *threo* over *erythro* (Eq.S4).[1] We will then proceed to establish *dr* as a function of er_2 and er_3, giving us a possible method of directly calculating *dr*. Here *A*, *B*, *C*, and *D* represent the concentrations of these structures, i.e. the total speciation we are seeking.

$$er = \frac{R}{S} \qquad \text{Eq.S1}$$

$$er_2 = \frac{x}{y} = \frac{A+B}{C+D} \qquad \text{Eq.S2}$$

$$er_3 = \frac{u}{v} = \frac{A+C}{B+D} \qquad \text{Eq. S3}$$

$$dr = \frac{threo}{erythro} = \frac{A+D}{B+C} \qquad \text{Eq. S4}$$

Let us also observe that:

$$\frac{dr}{dr} = \frac{1}{dr} \times dr = \frac{(B+C)(A+D)}{(A+D)(B+C)} = 1 \qquad \text{Eq. S5}$$

In order to establish a relationship between *dr*, er_2 and er_3, we used an algebraic manipulation involving a fractional expression, i.e. $((er_2 \times er_3) + 1)/(er_2 + er_3)$ and evaluate it by substituting the definitions from Eqs. 2 and 3, and multiplying the terms out to get:

$$f(er_2, er_3) = \frac{(er_2 \times er_3) + 1}{(er_2 + er_3)} \qquad \text{Eq.S6a}$$

$$= \frac{xu + yv}{xv + yu} \qquad \text{Eq.S6b}$$

$$= \frac{A^2 + AC + AB + 2BC + BD + CD + D^2}{B^2 + BD + AB + 2AD + CA + CD + C^2} \qquad \text{Eq.S6c}$$

We can multiply the final expression shown above by $\frac{dr}{dr} = 1$ (shown in Eq.S5) using the property of the multiplicative identity to keep the value the same. We can then expand the first two expressions in both the numerator and the denominator to find like terms:

$$= \frac{(A^2 + AC + AB + 2BC + BD + CD + D^2)(B+C)(A+D)}{(B^2 + BD + AB + 2AD + CA + CD + C^2)(A+D)(B+C)} \qquad \text{Eq.S7a}$$

[1]Enantiomeric and diastereomeric ratios are interchangeable with enantiomeric excesses and diastereomeric excesses, respectively. Ratios were chosen over excesses to simplify the derivation of Eq. 6a.

$$= \frac{(A^2C + AC^2 + ABC + BCD + C^2D + CD^2 + 2BC^2 + A^2B + AB^2 + ABC + BCD + B^2D + BD^2 + 2B^2C)(A + D)}{(A^2C + AC^2 + ABD + ACD + C^2D + CD^2 + 2AD^2 + A^2B + AB^2 + ABD + ACD + B^2D + BD^2 + 2A^2D)(B + C)} \quad \text{Eq. S7b}$$

$$= \frac{(A^2C + AC^2 + C^2D + CD^2 + A^2B + AB^2 + B^2D + BD^2 + 2ABC + 2BCD + 2BC^2 + 2B^2C)(A + D)}{(A^2C + AC^2 + C^2D + CD^2 + A^2B + AB^2 + B^2D + BD^2 + 2ABD + 2ACD + AD^2 + 2A^2D)(B + C)} \quad \text{Eq.S7c}$$

Now, we can define the variable $Z = (A^2C + AC^2 + C^2D + CD^2 + A^2B + AB^2 + B^2D + BD^2)$, which corresponds with the first eight terms in both the numerator and the denominator of the of Eq.S7c. Hence, we can substitute in Z and rewrite the above as:

$$= \frac{(Z + 2ABC + 2BCD + 2BC^2 + 2B^2C)[(A + D)]}{(Z + 2ABD + 2ACD + AD^2 + 2A^2D)[(B + C)]} \quad \text{Eq.S8}$$

We now see that the portion of expression in Eq.S8 put into brackets on the right corresponds exactly with our definition of *dr* as shown in Eq.S4. Thus, we can also pull out the like terms in the latter halves of the numerator and denominator ($2BC$ and $2AD$, respectively), and simplify the expression as follows:

$$= \frac{(Z + 2ABC + 2BCD + 2BC^2 + 2B^2C)}{(Z + 2ABD + 2ACD + AD^2 + 2A^2D)} \times dr \quad \text{Eq.S9a}$$

$$= \frac{Z + BC \cdot 2(A + B + C + D)}{Z + AD \cdot 2(A + B + C + D)} \times dr \quad \text{Eq.S9b}$$

It is now evident that the only terms that differ in this fraction are BC and AD. Let us assume that $BC = AD = X$, and substitute this value into the equation, as we did when evaluating Eq.S7c to get Eq.S8. Then:

$$= \frac{Z + X \cdot 2(A + B + C + D)}{Z + X \cdot 2(A + B + C + D)} \times dr \quad \text{Eq.S10a}$$

$$= (1)\, dr \quad \text{Eq.S10b}$$

So, with our above assumption intact, we have explicitly shown that Eq.S6a equals *dr*:

$$f(er_2, er_3) = \frac{(er_2 \times er_3) + 1}{(er_2 + er_3)} = dr$$

Therefore, we now know that only if $BC = AD$, can we directly calculate *dr* from a function of er_2 and er_3 (Eq.S6a).

II. Three-dimensional scatterplots of generated er_2 and er_3 values and the deviation of the calculated *dr* with the actual *dr* value using the derivation from **SI**.

The purple data points represent percent compositions that do not satisfy the special circumstance, whereas the cyan data points represent percent compositions that satisfy the special circumstance.

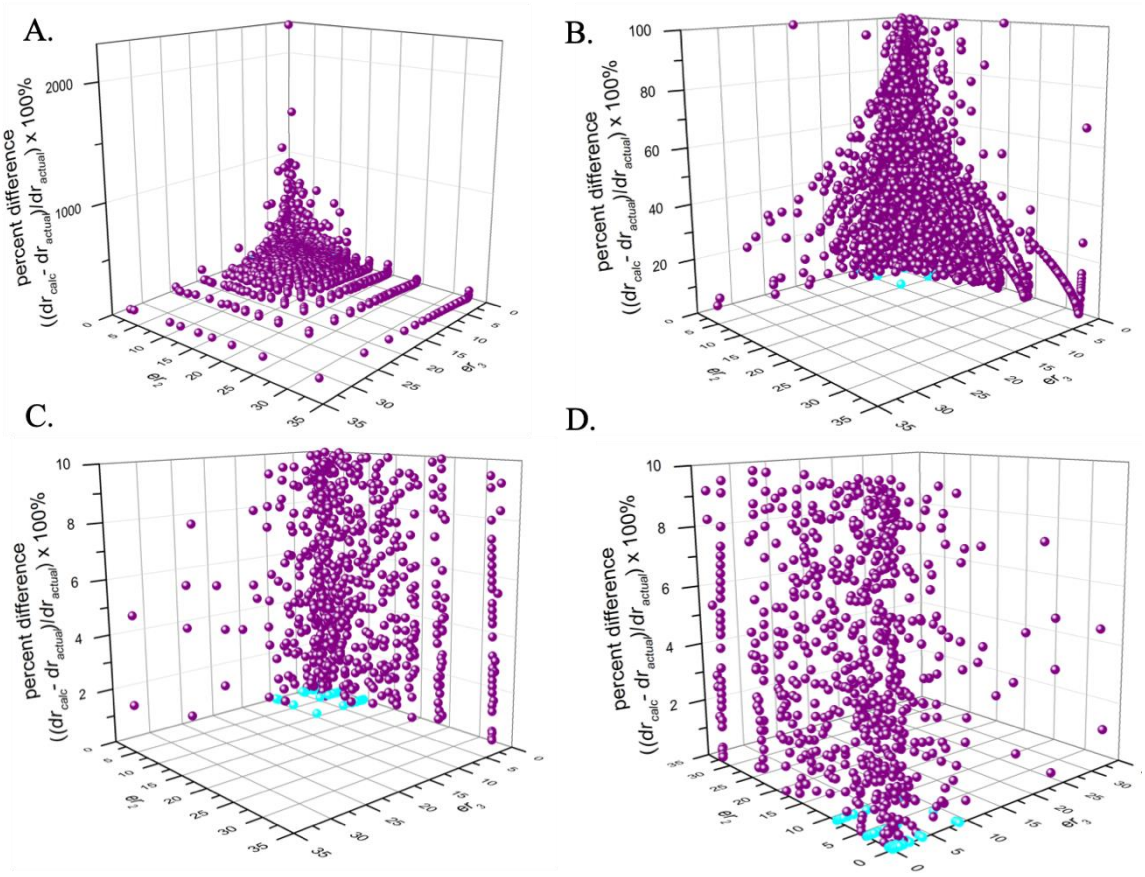

Figure S1A. Three-dimensional scatterplot of the generated er_2 and er_3 values (x- and y-axis, respectively) and the deviation of the calculated *dr* with the actual *dr* value using Eq. 6.

S1B. Three-dimensional scatterplot of the generated er_2 and er_3 values (x- and y-axis, respectively) and the deviation of the calculated *dr* with the actual *dr* value using Eq. 6. blown up to show the percent difference of derived *dr* values to actual *dr* values that are less than 100%.

S1C. Three-dimensional scatterplot of the generated er_2 and er_3 values (x- and y-axis, respectively) and the deviation of the calculated *dr* with the actual *dr* value using Eq. 6. blown up to show the percent difference of derived *dr* values to actual *dr* values that are less than 10%.

S1D. Alternative view of Three-dimensional scatterplot of the generated er_2 and er_3 values (x- and y-axis, respectively) and the deviation of the calculated *dr* with the actual *dr* value using Eq. 6. blown up to show the percent difference of derived *dr* values to actual *dr* values that are less than 10%.

III. Derivation for the complete speciation for a scalemic mixture of four stereoisomers using enantiomeric excess (*ee*) values at individual stereocenters and a diastereomeric excess value (*de*).

We now turn to an analysis on how to accomplish a complete speciation of all four stereoisomers if the enantioenrichment at the two separate stereocenters and the diastereomeric excess (*de*) is known.[2] We begin by defining the *ee* as the enrichment of *R* relative to *S* (Eqs.S11 and S12) and the *de* as *erythro* relative to *threo* (Eq.S13a).

$$ee_2 = \%\,2R - \%\,2S \qquad \text{Eq.S11a}$$

$$ee_2 = \frac{(2R,3R) + (2R,3S) - (2S,3R) - (2S,3S)}{(2R,3R) + (2R,3S) + (2S,3R) + (2S,3S)}\ X\ 100\% \qquad \text{Eq.S11b}$$

$$ee_2 = \left(\frac{A + B - C - D}{A + B + C + D}\right) X\ 100\% \qquad \text{Eq.S11c}$$

$$ee_3 = \%\,3R - \%\,3S \qquad \text{Eq.S12a}$$

$$ee_3 = \frac{(2R,3R) + (2S,3R) - (2R,3S) - (2S,3S)}{(2R,3R) + (2R,3S) + (2S,3R) + (2S,3S)}\ X\ 100\% \qquad \text{Eq.S12b}$$

$$ee_3 = \left(\frac{A + C - B - D}{A + B + C + D}\right) X\ 100\% \qquad \text{Eq.S12c}$$

$$de = \%\ threo - \%\ erythro \qquad \text{Eq.S13a}$$

$$de = \frac{(2R,3R) + (2S,3S) - (2R,3S) - (2S,3R)}{(2R,3R) + (2R,3S) + (2S,3R) + (2S,3S)} X\ 100\% \qquad \text{Eq.S13b}$$

$$de = \left(\frac{A + D - B - C}{A + B + C + D}\right) X\ 100\% \qquad \text{Eq.S13c}$$

First, we must express the percent composition of stereoisomers possessing *2R*-, *2S*-, *3R*-, and *3S*-handedness using the ee_2, and ee_3 values (Eqs.S14-S17). We can also express the percent composition of *erythro* and *threo* stereoisomers using the *de* values (Eqs.S18 and S19).

$$A + B = \left(\frac{ee_2}{2}\right) + 50 \qquad \text{Eq.S14}$$

$$C + D = \left(\frac{-ee_2}{2}\right) + 50 \qquad \text{Eq.S15}$$

$$A + C = \left(\frac{ee_3}{2}\right) + 50 \qquad \text{Eq.S16}$$

$$B + D = \left(\frac{-ee_3}{2}\right) + 50 \qquad \text{Eq.S17}$$

[2] Excesses were chosen over ratios to simplify the percent composition derivation.

$$A + D = \left(\frac{de}{2}\right) + 50 \qquad \text{Eq.S18}$$

$$B + C = \left(\frac{-de}{2}\right) + 50 \qquad \text{Eq.S19}$$

We can now express the components *B*, *C*, and *D* in terms of *A* (Eqs.S20-S22).

$$B = \left(\frac{ee_2}{2}\right) + 50 - A \qquad \text{Eq.S20}$$

$$C = \left(\frac{ee_3}{2}\right) + 50 - A \qquad \text{Eq.S21}$$

$$D = \left(\frac{de}{2}\right) + 50 - A \qquad \text{Eq.S22}$$

Using the mass balance in terms of percent composition (Eq. S23) and substituting Eqs.S20-S22 for *B*, *C*, and *D*, respectively, we arrive at Eq.S24a. By combining like terms (Eqs.S24b, S24c) and expressing *A* in terms of ee_1, ee_2 and *de*, we can now use Eqs.S20-S22 to determine the complete speciation.

$$A + B + C + D = 100 \qquad \text{Eq.S23}$$

$$A + \left[\left(\frac{ee_2}{2}\right) + 50 - A\right] + \left[\left(\frac{ee_3}{2}\right) + 50 - A\right] + \left[\left(\frac{de}{2}\right) + 50 - A\right] = 100 \qquad \text{Eq.S24a}$$

$$A + \frac{ee_2}{2} + \frac{ee_3}{2} + \frac{de}{2} + 150 - 3A = 100 \qquad \text{Eq.S24b}$$

$$\frac{ee_2}{2} + \frac{ee_3}{2} + \frac{de}{2} - 2A = -50 \qquad \text{Eq.S24c}$$

$$A = \frac{1}{4}ee_2 + \frac{1}{4}ee_3 + \frac{1}{4}de + 25 \qquad \text{Eq.S24d}$$

Thus, we have shown that a four component stereoisomeric mixture can be fully characterized from readily measurable experimental parameters (ee_2, ee_3 and *de* values).

IV. Validation of complete speciation mathematics via a few randomly generated percent compositions.

Table S1. Validation of Speciation Math

(2R,3R)	*(2R,3S)*	*(2S,3R)*	*(2S,3S)*					
A	B	C	D	Total	*ee2*	*ee3*	*de*	*(ee2 + ee3 +de) /4*
15	35	40	10	100	0	10	-50	-10
15	30	20	35	100	-10	-30	0	-10
10	70	10	10	100	60	-60	-60	-15
20	10	6	64	100	-40	-48	68	-5