Pergamon

Tetrahedron: Asymmetry, Vol. 8, No. 8, pp. 1153–1156, 1997
© 1997 Elsevier Science Ltd. All rights reserved.
Printed in Great Britain
PII: S0957-4166(97)00112-2
0957-4166/97 $17.00 + 0.00

Kinetic resolutions: evaluation of a one dimensional E value calculation method using a computer and statistical software

Milton J. Zmijewski,[a,*] Gary Sullivan,[a] Rose Persichetti [b] and Jim Lalonde [b]

[a] Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, USA
[b] Altus Biologics, Inc., 40 Allston Street, Cambridge, MA 02139-4211, USA

Abstract: A recently published curve fitting method of determining E values was put to the test using a computer and statistical software. Using idealized data and curves, the methodology worked reasonably well. Confidence limits generated by the ideal data were still large. However, when real data were employed, this methodology was not satisfactory in generating reliable E values due to inherent errors in the progress curves generated. This approach does not appear to be suitable for determining E values from one dimensional progress curves even when powerful statistical software and a computer are used to evaluate the data. © 1997 Elsevier Science Ltd

Since the publication of the Sih method for E value or enantiomeric ratio calculations[1] of enzymatic kinetic resolutions, several reports of other methods of doing this same calculation have appeared.[2–5] The Sih method requires the use of the conversion (c) and ee (enantiomeric excess) of the product or substrate. More recently, Rakels *et al.*[2] have developed a method that only used the ee of the product and substrate and is independent of the amount of starting material converted. An easy to use computer program has been published on the internet (http://www-orgc.tu-graz.ac.at/) that allows one to use either method in E value calculations (for both mac and pc platforms). The critical E value in both cases is 15. Values below 15 are considered low and give low level resolutions. Between 15 and 30, resolutions are acceptable. Numbers above 30, of course, are the most desirable and give the highest yield and ee of product or substrate.

A simple curve fitting approach has been sought to allow the time course of the reaction to serve as the starting point for E value calculation.[3–5] Such an approach would be extremely useful in cases where a good chiral method to determine ee is not available or in situations where many enzymes and/or substrates have to be screened. Errors in using progress curves alone are numerous and have been described previously.[4] Of primary concern are enzyme inhibition and deactivation, side reactions, and changes in the system as the reaction progresses. A complicated method for calculating E values from progress curves has been described[3] but was only evaluated for low E value systems.

A more simplified method has recently been published[5] and is the subject of this article. As with the previous method, it showed data for low E value resolutions only. This method uses equation (1) and simple algebra to solve for E. In this equation, **S/So** is the portion of substrate remaining, **k** is the rate constant for the fast reacting enantiomer, and **E** is the ratio of the specificity constants of the fast reacting isomer divided by that of the slower reacting isomer. The authors claim that only two time points of the progression curve are needed to estimate E if the substrate concentration is known at t1 and t2=2t1. The problem is that different pairs of points will provide different solutions for E. This variability was seen in tests of this method (Sullivan unpublished results).

$$S/S_0 = 0.5(e^{-k \cdot time} + e^{-k \cdot time/E}) \qquad \text{Equation (1)}$$

* Corresponding author. Email: zmijew@lilly.com

Table 1. E Values from idealized data# and calculated using JMP®

E value from Idealized Data	E value calculated using JMP®	Lower Confidence Limit	Upper Confidence Limit
5	5.23	2.8	12.8
20	15.3	12.7	18.7
50	47	31	80.6

#All entries obtained from Ref.[3] Data from Ref.[3] was interpolated from the theoretical graphs in the publication.

Table 2. E Values calculated using the Sih and JMP® methods* and experimental data

E value from experimental data[1]	E value calculated using JMP® method	Lower Confidence Limit	Upper Confidence Limit
1*	3	2	4
6.8#	4.52	4.2	4.9
30^	91	57	158
97.5+	296	275	319

*Data from Ref.[5] #See note in Ref.[9] ^See note in Ref.[10] +See note in Ref.[11]

We felt that the best approach to using this methodology is to estimate E using all the data and non-linear modeling.[6] The primary advantages of doing so are that all the data are used simultaneously and the estimated value of E is accompanied by an estimated standard deviation (or standard error) from which a 95% confidence interval can be determined. The perceived disadvantage may be the complexity of the method. However, a statistical software package like JMP®[7] will easily estimate the unknown parameters **k** and **E** from equation (1) with as few as four points and with minimal effort on the part of the scientist. JMP® uses the Gauss–Newton method of non-linear modeling to estimate the parameters of equation (1). This iterative method requires starting values for the unknown parameters, **k** and **E**. The method then predicts the substrate concentration at the starting values of the parameters. The substrate concentration residuals (the differences between the observed and predicted substrate concentrations) are determined and subsequently minimized using the method of least squares. At each step of the procedure, the estimates of **k** and **E** are adjusted to produce a maximum reduction in the residuals. The procedure is finished or converges once the residuals fall below a preset limit. Confidence limits can also be determined once the parameters are calculated.

The idealized results from Rakels *et al.*[3] are shown in Table 1. Based on these results, our curve fitting analysis of the interpolated data from the paper would seem to give a reasonably good approximation of E value. These data, however, are idealized and do not have any of the potential problems that can occur in real kinetic resolutions. Even though the curves are generated from idealized data, there is still a large spread in the confidence limits of **E** using our curve fitting method.

Using data from Lu *et al.*[5] and our own data (Table 2) generated with soluble enzymes and enzymes in their CLEC® form, we found real differences between the Sih method and this non-linear modeling to find the E value. Using all the data from Lu *et al.*[5] and non-linear modeling (entry 1, Table 2), we obtained a higher E value as opposed to using two data points and algebra. In another case (entry 2, Table 2), the non-linear method gave an E value below that calculated by the Sih approach. In entry 3 and 4 of Table 2, the E value was much higher using this method than using the Sih method. At best, the one dimensional non-linear modeling method using JMP® gives an approximation of the E value. Data not published here on proprietary compounds gave similar results.[8]

The authors recommend that the JMP® curve fitting method for E value calculations be used in only two situations; first, in reactions where it is known *a priori* that the enzyme kinetics follow ideal conditions. These reactions result in E values that compare favorably with the Sih method (Table 1). It is likely that this situation will be rare but could arise. Second, it may be feasible to use this method to prioritize a series of related compounds being resolved with one enzyme. Results on proprietary

compounds[8] indicate that the curve fitting method can give similar E value rankings of compounds to the Sih method. The compounds must be close in structure and the same enzyme used with the entire series of compounds. This situation is often encountered in industrial process chemistry development where a critical lead intermediate might change several times during the course of synthetic route selection. Rather than developing a new chiral method for each compound, the curve fitting method could be employed as a pre-screen. Confirmation of the absolute E value would have to be done subsequently using the standard methods.

Based on these results, the authors do not feel that the Sih method or the Rakels modification should be replaced with this one dimensional method except in certain restricted situations. If used, the method should be used cautiously since it can not discount the normal interferences that occur when following reaction progress curves. Using a statistical program such as JMP® to perform the curve fitting is imperative to using the non-linear method. Two data points will not give a curved line with reasonable confidence limits. Since progress curves are full of inherent errors anyway, a two point method only invites additional error introduction. Where possible, confirmation should ultimately be obtained using a second dimension in evaluating the data. The methods of Sih or Rakels, while known to be inaccurate at high E values, are preferred over this non-linear method calculated using JMP®. This is because both of these well established methods require two dimensions of the enzymatic resolution to be measured.

References

1. Chen, C. S; Fujimoto, Y.; Girdaukas, G.; Sih, C. J. *J. Amer. Chem. Soc.* **1982**, *104*, 7294. $E=[\ln[1-c(1+ee_P)]]/[\ln[1-c(1-ee_P)]]$ or $[\ln[(1-c)(1-ee_S)]]/[\ln[(1-c)(1+ee_S)]]$. E is the enantiomeric ratio; c is the conversion and ee is the enantiomeric excess of the substrate (S) or product (P).
2. Rakels, J. L. L.; Straathof, A. J. J.; Heijnen, J. J. *Enzyme Microb. Technol.* **1993**, *15*, 1051. $E=\ln[(1-ee_S)/(1+(ee_S/ee_P))]/\ln[(1+ee_S)/1+(ee_S/ee_P))]$.
3. Rakels, J. L. L.; Romein, B.; Straathof, A. J. J.; Heijnen, J. J. *Biotech. Bioeng.* **1994**, *43*, 411.
4. Fourneron, J. D.; Combemorel, A.; Buc, J.; Pieroni, G. *Tetrahedron Lett.* **1992**, *33*, 2469.
5. Lu, Y.; Zhao, X.; Chen, Z. N. *Tetrahedron Asymm.* **1995**, *6*, 1093.
6. Draper, N. R.; Smith, H. *Applied Regression Analysis* Wiley, New York, **1981**.
7. *JMP® Statistics and Graphics Guide* Version 3.1, SAS Institute, Inc., **1995**. For this application, equation (1) is entered as the 'formula' for one of the columns of the spread sheet. This becomes the F(x) column. 'Time' and the '% substrate remaining' are then entered into separate data columns. The non-linear curve fitting application is selected, the F(x) column is entered into the 'X' field and the '% substrate remaining' is entered into the 'Y' field. The 'go' button is selected and the iterative process begins. The method is little more complicated then graphing data in a simple graphing program.
8. Zmijewski, M. J.; Briggs, B. unpublished data.
9. Methyl bromopropionate (0.736g, 4.4 mmol) in 50 ml of 20 mM pH 5 phosphate buffer cooled to 5°C. *Candida rugosa* esterase (20 mg/ml) was added and the pH maintained with base addition. Progress curve determined by amount of based added. Reaction was complete after 120 minutes. Chiral GC conditons: Cyclodex B 25 m capillary GC column; He flow 1 ml/min; initial temperature=50°C, gradient rate=5°C/min, final temp.=130°C. Retention times for the methyl ester enantiomers were 12.6 and 13.25 min.
10. (R,S)-Sulcatol (50 mg) and vinyl acetate (100 μl) were added to 5 ml of toluene. ChiroCLEC-PC (Altus Biologics) (5.2 mg) was added to start the reaction. Progress curves were generated by following the reaction using chiral GC. Chiral GC conditions: Cyclodex B capillary GC 25 m column; He flow=1 ml/min; initial temperature=90°C, gradient=5°C/min, final temp.=130°C. Retention times; (S)-sulcatol, 12.5 min; (R)-sulcatol, 12.32 min; (S)-sulcatol acetate, 14.43 min;

(R)-sulcatol acetate, 15.12 min. The ee of the remaining starting material and product alcohol was 99% and 69.1% respectively at 59% conversion.

11. ChiroCLEC-CR (Altus Biologics) (2 mg) was added to a solution of (±)-trans-2-methyl cyclohexanol (0.2 mmol) in 1 ml of toluene containing 1 μl water. Vinyl butyrate (0.2 mmol) was added after a 1 minute stir. Progress curves at 25°C were generated using capillary GC. GC conditions: DB1701 column; He flow rate=25 cm/sec; initial temp.=60°C, gradient=10°C/min, final temp.=170°C. Retention times: alcohol, 4.6 min; ester, 8.43 min. Chiral GC was similar to above method.

(Received in USA 3 February 1997)