JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Model Building Using Linear Free Energy Relationship Parameters—Eliminating Calibration Curves for Optical Analysis of Enantiomeric Excess

Chung-Yon Lin, Stephanie Lim, and Eric V. Anslyn*

Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States

Supporting Information

ABSTRACT: Linear free energy relationship (LFER) parameters are routinely used to parametrize physicochemical effects while investigating reaction mechanisms. In this Communication, we describe an alternate application for LFERs: training sets for model building in an analytical application. In this study, the sterics, quantified by Charton parameters (Δv), of nine secondary chiral alcohol analytes were correlated to the circular dichroism output from a chiral alcohol optical sensor. To test the validity of the model, the correlative linear model was applied to determine the enantiomeric excess of samples of two alcohols without *a priori* knowledge of a calibration curve. The error in this method was comparable to those of previous experimental methods (<5%).

Physical organic chemists employ many tools to elucidate reaction mechanisms. Within this suite of tools, linear free energy relationships (LFERs) utilize a set of parameters (e.g., Taft or Hammett parameters) to quantify a physicochemical substituent effect (e.g., steric or electronic) in a reference system.[1,2] Establishing correlations between these effect-specific parameters with measurements of relative free energy in a new system generates a measurement of the sensitivity to the substituent effects (slope of the correlation). The interpretation of this sensitivity and of the correlation goodness-of-fit offers insight into a reaction mechanism.[3−8] However, in addition to mechanistic insights, LFERs can also be used to quantify a physicochemical effect for other purposes.[7,9−12] For example, in the closely associated field of quantitative structure−activity relationships (QSARs), LFERs are routinely employed alongside computationally derived parameters to build predictive models for drug interactions.[13−19]

Recently, Mayr established correlations of nucleophiles (Nuc) and electrophiles (El) that can be used to predict the reactivity of Nuc−El combinations in a variety of reactions.[20−24] In fact, dating back to the early work of Ritchie[25] and Kane-Maquire,[26] LFERs have been recognized for their potential, not merely to analyze trends, but also to predict reaction rates and thermodynamics. Recent work by Sigman extended these ideas, demonstrating that multiparameter models, with both LFERs and parameters derived in silico, can be used to optimize catalytic reaction conditions.[5,8−10,15,27−30] We wondered whether LFERs could model optical phenomena to generate predictive models, as has been done with reactivity.

Our group has developed optical assays for enantiomeric excess (ee) that involve molecular recognition or multicomponent assemblies.[6,7,31] Circular dichroism (CD) signal that arise upon binding of a chiral analyte are correlated to the sample's ee values through experimentally derived calibration curves. To date, efforts to generate structure−activity predictive models have been primarily focused upon UV and fluorescence spectra of small organic molecules or assembly.[17,32−34] Doing so with the CD spectra generated from our chiral sensing systems would remove the necessity to generate a calibration curve for every unique chiral analyte prior to analysis.

We reasoned that such modeling would be successful because for each optical-based ee sensor the absorptivity and the HOMO/LUMO gaps do not change significantly and because the CD signal modulations are purely dependent on chemical attributes of the analytes. Thus, the correlation of analyte structural descriptors (LFERs) with the optical output should elucidate the physicochemical origin behind the signal modulation, thereby building predictive models for analytes not involved in the training set.

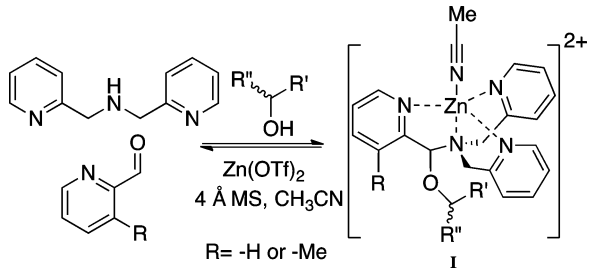
In several of our ee sensor studies, we demonstrated correlations between analyte sterics and the diastereomeric forms of enantiomeric complexes that result in the different CD signals.[7,11,12]The analyte sterics were parametrized with LFERs such as those of Taft and Charton. Taft parameters were developed by measuring rates of acid-catalyzed hydrolysis of substituted methyl esters (RCOOMe), where the changes in rate were a result of steric interactions between the R group and the nucleophile.[1]Charton observed a correlation between Taft parameters and the van der Waals radii of symmetric substituents and adjusted the values of the Taft parameters to agree with their corresponding van der Waals radii.[35,36]

To explore whether predictive correlations could be created for CD ellipticities, we examined the multicomponent assembly shown in Scheme 1. In this assembly, the steric differences in the R groups on the stereocenter influence the populations of the diastereiomers of **1** and set a helical twist into the pyridines, giving rise to exciton-coupled CD. We recently reported that a 3-methyl variant of 2-formylpyridine produced a larger CD dynamic range and a concomitant lower error in ee determination.[12]

Received: April 16, 2016
Published: June 15, 2016

DOI: 10.1021/jacs.6b03928
J. Am. Chem. Soc. 2016, 138, 8045−8047

Scheme 1. Multicomponent Assembly Formed with Variants of Pyridine-2-Carbaldehyde, Di-2-picolylamine, Zn(II), and a Chiral Alcohol; The CD-Active Hemiaminal Ether Species Is Shown as 1



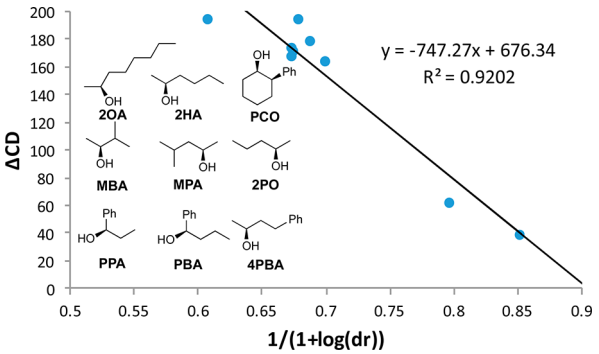


Figure 1. The nine alcohols used in the data set and a plot showing the correlation between the assembly diastereomeric ratio (dr) and ΔCD. The unit of ΔCD is mdeg.

Table 1. ΔCD and dr Values of the Nine Alcohols Used in the Correlative Analysis

alcohol	ΔCD (mdeg)	$[1 + \log(dr)]^{-1}$
2OA	173.26	0.68
2HA	193.50	0.68
2PO	163.02	0.70
MPA	193.53	0.61
PCO	170.33	0.68
MBA	167.06	0.68
PBA	61.80	0.80
PPA	38.36	0.85
4PBA	177.82	0.69

In the present study, we used the 3-methyl variant of the CD-active assembly with a training set of nine alcohols (Figure 1). As expected, a correlative trend between the CD signals and the diastereomeric ratio (dr) was found using a previously derived equation (Figure 1 and Table 1).[6] Given this positive result, we set out to derive a correlation between the CD signals and structural parameters, such as Taft and Charton parameters, of the chiral alcohol analytes. In an attempt to incorporate computationally generated parameters, we included the use of sterimol, which provides parameters that describe the fine-grained three-dimensional structure.[37] Sterimol can discern slight substituent steric differences whereas the experimentally based parameters cannot.

With the large pool of sterimol parameters generated for all of the chiral alcohols, the "leaps" package in R software was used to perform a variable selection algorithm. The algorithm considers all possible regression models with the given set of variables and uses adjusted R^2 and the Bayes information criterion as selection criteria. To our dismay, the sterimol parameters were not as strongly correlated to the assembly dr values as the simpler Charton parameter model was (see the Supporting Information). We hypothesized this was due to the thermodynamic nature of the sensor–analyte assembly, for which numerous conformations of the substituents on the hemiaminal ether in **1** (R′ and R″ in Scheme 1) are possible. Although sterimol provides a more detailed steric characterization, the descriptors are for a static substituent. In contrast, the partially experimentally derived Charton parameters better account for conformational effects. Thus, the Charton parameters quantifying the steric differences of chiral alcohol substituents (Δv) were correlated to the assembly dr (R^2 = 0.83; Figure 2).[35,36,38−40] At this point, we have a correlative model relating the substituent difference, quantified by Δv, to the dr that predicts the sensor CD output.

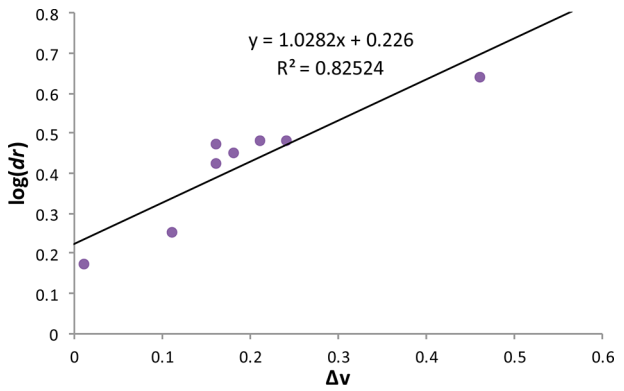


Figure 2. LFER correlating the difference in steric size for substituents on the stereocenter of a secondary alcohol to the assembly log(dr) value.

This model was used to generate theoretical calibration curves for two alcohols that were not in the training set: 2-butanol and 1-phenylethanol. The difference in Charton parameter for non-hydrogen groups were used to predict the CD values of enantiopure (100 and −100 ee) sample from the aforementioned model. Excitingly, the calibration curves from the predictive model resembled those from experimental results (Figure 3). The two test alcohols were chosen to have substituents within the range of steric sizes taken into account by the model. With the calibration curves from the predictive model, unknown samples were determined at 2.27% average error (Table 2), which is within the margin of error achieved by experimental calibration curves.[11,41] It is worth noting that the two test alcohols were chosen to be within the range of the

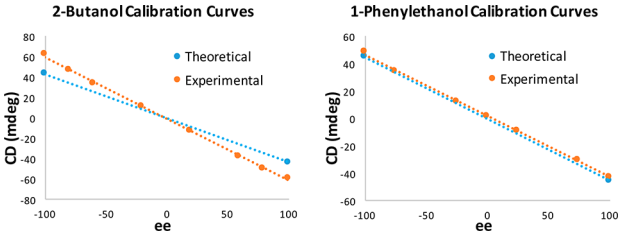


Figure 3. Overlay of calibration curves for 2-butanol (left) and 1-phenylethanol (right). The theoretical curves were generated with the predictive model, while the experimental calibration curves were generated with experimental values.

DOI: 10.1021/jacs.6b03928
J. Am. Chem. Soc. 2016, 138, 8045−8047

Table 2. Determination of the Errors in ee for 2-Butanol and 1-Phenylethanol Using Calibration Curves Generated by a Predictive Model

2-Butanol			
ee	CD	calcd ee	% error
−60.00	27.16	−62.73	2.73
40.00	−20.71	47.84	7.84
−20.00	9.07	−20.94	0.94
			avg: 3.83
1-Phenylethanol			
ee	CD	calcd ee	% error
50.00	−19.60	43.56	6.44
10.00	−2.40	5.33	−4.67
−40.00	18.15	−40.33	0.33
			avg: 0.70
			overall avg: 2.27

model established above. For alcohols with other functionalities that are not alkyl or phenyl, it is expected that additional data would be required to extend the range of the current model.

In summary, we found that an LFER removes the necessity to experimentally generate a calibration curve for an optical sensor, as steric parameters accurately predicted a chiroptical response. To our knowledge, this is the first such application of LFERs, and we are currently applying this approach to our other CD assays. Importantly, the lesson derived from our studies is to consider LFERs as model-building data to be used in statistical analyses and analytical chemistry. Admittedly, this insight may seem obvious, particularly after nearly 80 years history of LFERs, but it could reinvigorate the usefulness of this classic physical organic technique and bring to LFERs increased power and utility.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b03928.

Variable selection algorithm and data output (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*anslyn@austin.utexas.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The project was financially supported by the NIH (R01-GM077437) and the Welch Regents Chair (F-0046).

■ REFERENCES

(1) Taft, R. W. *J. Am. Chem. Soc.* **1953**, *75*, 4538−4539.
(2) Hammett, L. P. *J. Am. Chem. Soc.* **1937**, *59*, 96−103.
(3) Hammett, L. P. *Chem. Rev.* **1935**, *17*, 125−136.
(4) Akhani, R. K.; Moore, M. I.; Pribyl, J. G.; Wiskur, S. L. *J. Org. Chem.* **2014**, *79*, 2384−2396.
(5) Gustafson, J. L.; Sigman, M. S.; Miller, S. J. *Org. Lett.* **2010**, *12*, 2794−2797.
(6) You, L.; Berman, J. S.; Lucksanawichien, A.; Anslyn, E. V. *J. Am. Chem. Soc.* **2012**, *134*, 7126−7134.
(7) Joyce, L. a; Maynor, M. S.; Dragna, J. M.; da Cruz, G. M.; Lynch, V. M.; Canary, J. W.; Anslyn, E. V. *J. Am. Chem. Soc.* **2011**, *133*, 13746−13752.
(8) Harper, K. C.; Sigman, M. S. *J. Org. Chem.* **2013**, *78*, 2813−2818.
(9) Milo, A.; Bess, E. N.; Sigman, M. S. *Nature* **2014**, *507*, 210−214.
(10) Neel, A. J.; Milo, A.; Sigman, M. S.; Toste, F. D. *J. Am. Chem. Soc.* **2016**, *138*, 3863−3875.
(11) You, L.; Berman, J. S.; Anslyn, E. V. *Nat. Chem.* **2011**, *3*, 943−948.
(12) Lin, C.-Y.; Giuliano, M. W.; Ellis, B. D.; Miller, S. J.; Anslyn, E. V. *Chem. Sci.* **2016**, DOI: 10.1039/C5SC04629G.
(13) Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J. C.; Sheridan, R. P.; Feuston, B. P. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947−1958.
(14) Fujita, T.; Winkler, D. A. *J. Chem. Inf. Model.* **2016**, *56*, 269−274.
(15) Harper, K. C.; Bess, E. N.; Sigman, M. S. *Nat. Chem.* **2012**, *4*, 366−374.
(16) Huang, H.; Zong, H.; Bian, G.; Yue, H.; Song, L. *J. Org. Chem.* **2014**, *79*, 9455−9464.
(17) Katritzky, A. R.; Kuanar, M.; Slavov, S.; Hall, C. D.; Karelson, M.; Kahn, I.; Dobchev, D. A. *Chem. Rev.* **2010**, *110*, 5714−5789.
(18) Denmark, S. E.; Gould, N. D.; Wolf, L. M. *J. Org. Chem.* **2011**, *76*, 4337−4357.
(19) Huang, H.; Zong, H.; Bian, G.; Song, L. *J. Org. Chem.* **2012**, *77*, 10427−10434.
(20) Mayr, H.; Patz, M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 938−957.
(21) Appel, R.; Chelli, S.; Tokuyasu, T.; Troshin, K.; Mayr, H. *J. Am. Chem. Soc.* **2013**, *135*, 6579−6587.
(22) Guo, X.; Mayr, H. *J. Am. Chem. Soc.* **2014**, *136*, 11499−11512.
(23) Mayr, H.; Ammer, J.; Baidya, M.; Maji, B.; Nigst, T. A.; Ofial, A. R.; Singer, T. *J. Am. Chem. Soc.* **2015**, *137*, 2580−2599.
(24) Mayr, H.; Bug, T.; Gotta, M. F.; Hering, N.; Irrgang, B.; Janker, B.; Kempf, B.; Loos, R.; Ofial, A. R.; Remennikov, G.; Schimmel, H. *J. Am. Chem. Soc.* **2001**, *123*, 9500−9512.
(25) Ritchie, C. D. *Acc. Chem. Res.* **1972**, *5*, 348−354.
(26) Kane-Maguire, L. A. P.; Honig, E. D.; Sweigart, D. A. *Chem. Rev.* **1984**, *84*, 525−543.
(27) Bess, E. N.; Bischoff, A. J.; Sigman, M. S. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 14698−14703.
(28) Mei, T.-S.; Patel, H. H.; Sigman, M. S. *Nature* **2014**, *508*, 340−344.
(29) Harper, K. C.; Vilardi, S. C.; Sigman, M. S. *J. Am. Chem. Soc.* **2013**, *135*, 2482−2485.
(30) Bess, E. N.; Deluca, R. J.; Tindall, D. J.; Oderinde, M. S.; Roizen, J. L.; Du Bois, J.; Sigman, M. S. *J. Am. Chem. Soc.* **2014**, *136*, 5783−5789.
(31) Dragna, J. M.; Gade, A. M.; Tran, L.; Lynch, V. M.; Anslyn, E. V. *Chirality* **2015**, *27*, 294−298.
(32) Molnar, S. P.; King, J. W. *Int. J. Quantum Chem.* **1997**, *65*, 1047−1056.
(33) Williams, G. M.; Carr, R. A. E.; Congreve, M. S.; Kay, C.; McKeown, S. C.; Murray, P. J.; Scicinski, J. J.; Watson, S. P. *Angew. Chem., Int. Ed.* **2000**, *39*, 3293−3296.
(34) Pollock, J. B.; Schneider, G. L.; Cook, T. R.; Davies, A. S.; Stang, P. J. *J. Am. Chem. Soc.* **2013**, *135*, 13676−13679.
(35) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 1552−1556.
(36) Charton, M. *J. Org. Chem.* **1978**, *43*, 3995−4001.
(37) Verloop, A.; Tipker, J. *Pharmacochem. Libr.* **1987**, *10*, 97−102.
(38) Charton, M. *J. Am. Chem. Soc.* **1969**, *91*, 6649−6654.
(39) Charton, M. *J. Org. Chem.* **1976**, *41*, 2217−2220.
(40) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 3691−3693.
(41) Jo, H. H.; Lin, C.; Anslyn, E. V. *Acc. Chem. Res.* **2014**, *47*, 2212−2221.

DOI: 10.1021/jacs.6b03928
J. Am. Chem. Soc. 2016, 138, 8045−8047