JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Prediction of Catalyst and Substrate Performance in the Enantioselective Propargylation of Aliphatic Ketones by a Multidimensional Model of Steric Effects

Kaid C. Harper, Sarah C. Vilardi, and Matthew S. Sigman*

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: The effectiveness of a new asymmetric catalytic methodology is often weighed by the number of diverse substrates that undergo reaction with high enantioselectivity. Here we report a study that correlates substrate and ligand steric effects to enantioselectivity for the propargylation of aliphatic ketones. The mathematical model is shown to be highly predictive when applied to substrate/catalyst combinations outside the training set.

Although steric effects are ubiquitous in asymmetric catalysis, they remain difficult to quantify and define for precise application in catalyst design. Methods to measure steric effects and other effects in asymmetric catalysis have paralleled those developed for quantitative structure−activity relationships (QSAR) in drug design.[1] Typically, a QSAR study interrogates a single protein with a library of inhibitors to delineate the important substituent effect(s) responsible for the desired outcome.[2] The resultant mathematical models generated using QSAR afford a testable hypotheses about the key molecular interactions facilitating modern drug design.[2e]

In contrast to the specific protein−inhibitor interactions in medicinal chemistry, asymmetric catalysis is complicated because both the catalyst and the substrate are mutable and can be probed to determine the important structural features of each. Typically, catalysts have been examined independent of substrate or vice versa.[3] Indeed, the multifaceted interactions between both substrate and catalyst are generally ignored in examinations of asymmetric catalytic systems. Simultaneously evaluating the relationship between these crucial reaction variables is a goal of our ongoing program in understanding and predicting asymmetric catalytic reactions.[3g,4] Herein we present the results of the simultaneous and systematic evaluation of both catalyst and substrate steric effects on an enantioselective Nozaki−Hiyama−Kishi (NHK) ketone propargylation modeled with Sterimol parameters. Multidimensional modeling allows for accurate prediction of enantioselectivity for 41 independent substrate and catalyst combinations for aliphatic ketone propargylation.

Recently, we have reported several investigations of asymmetric catalytic reactions using linear free energy relationships where multiple steric and electronic components of the ligand structure were examined simultaneously.[4f−h] This multivariate examination of asymmetric catalytic reactions is appealing because it examines how different catalyst components might synergistically impact enantioselectivity. To date, these studies have focused exclusively on examining ligand components. Encouraged by the correlative and predictive models generated, we wanted to expand this approach to incorporate other common reaction parameters. We selected to probe what is perhaps the most impactful reaction consideration in asymmetric catalysis, the substrate, as we believed predicting and understanding substrate effects as a function of catalyst structure is at the core of how one develops and applies an asymmetric catalytic reaction.

Parameter Selection. Sterimol parameters vary from many common steric parameters because they describe a single substituent with three parameters, B_1, B_5, and L (Figure 1),

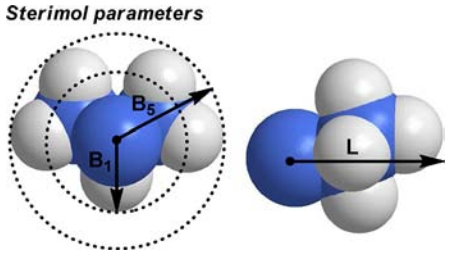


Figure 1. Parameterization of an isopropyl group using Verloop's Sterimol parameters.

instead of a single parameter.[5] The B_1 parameter describes a minimum width orthogonal to the primary bond, the B_5 parameter describes the maximum width along the same axis, and the L parameter is the length of the substituent along the primary bond. These separate parameters can generate more informative models than other simpler steric parameters but also increase the dimensionality of the data three-fold. In order to develop models using Sterimol parameters in combination with linear regression techniques, a larger data set is required than the nine-membered libraries previously utilized.[4f−h] Because the application of experimental design is fundamental to developing highly predictive models through multidimensional regression techniques, the catalyst substituents and ligand substituents need to be selected carefully using the statistical principles known as the design of experiments (DoE).[6]

Substrate and Catalyst Design Matrix. Our previous reports on the enantioselective propargylation of aliphatic

Received: January 7, 2013
Published: February 6, 2013

   dx.doi.org/10.1021/ja4001807 | J. Am. Chem. Soc. 2013, 135, 2482−2485

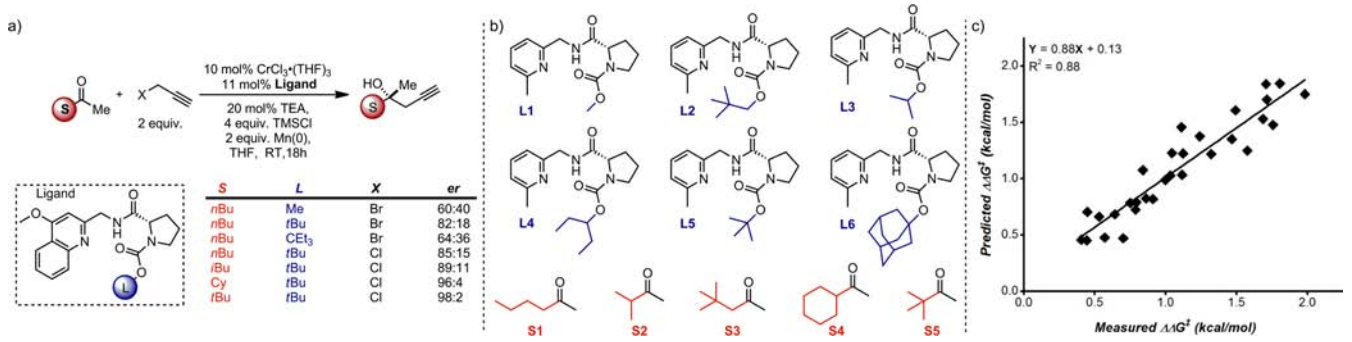


Figure 2. (a) Previous results showcasing sensitivity of both substrate and ligand steric effects. (b) Library design employing DoE principles. (c) Plot of experimental vs predicted enantioselectivity using the model in eq 2. A total of 30 unique ligand/substrate experiments were performed.

ketones using NHK conditions indicated that increased steric bulk on one side of the ketone led to higher enantioselectivity (Figure 2a).[4g] Coupling this sensitivity with the reported carbamate substituents effects on the ligand provided the foundation for the initial experimental design.[4c,d,f,g] Variation in steric bulk at the two substituents could be quantified using Sterimol parameters to develop a model capable of predicting enantioselectivity.

Our choice of substituents for the substrate and ligand was based on evenly distributing the Sterimol values for both according to DoE principles (Figure 2b). The Sterimol parameters present three steric dimensions to evenly span in order to produce an effective experimental design. Upon inspection of our previous efforts in enantioselective ketone propargylation, both the substrate and ligand have an apparent sensitivity to B_1 and B_5, allowing us to focus on an even distribution of the these parameters.[4g] The L parameter was also considered, but the collinear relationship between B_5 and L justified only ancillary attention.[4h] As we have previously reported, a precise experimental design is difficult to achieve using steric parameters because these parameters are not continuous.[4f] Accordingly, ligands **L1**−**L6** were chosen to have adequate and reasonably even variation for both B_1 and B_5 (see Supporting Information for details). Similarly, substrates **S1**−**S5** were selected.

Model Development. The target substrates are all commercially available. To facilitate the library synthesis, we developed a simplified methyl pyridine ligand as an attractive alternative to the previously reported quinoline-based ligand (Figure 1b).[4g] The advantages of using the methyl pyridine core are the rapid (in as few as three steps), modular, and scalable synthesis from commercial materials.

Propargylation of **S1**−**S5** using ligands **L1**−**L6** led to 30 unique observed enantioselectivities, each of which was replicated and averaged. Equation 1 is the base equation from which all models were derived. It contains the Sterimol parameters for the ligand substituent (B_{1L}, B_{5L}, and L_L) as well as the Sterimol parameters for the substrate substituent (B_{1S}, B_{5S}, and L_S). The base equation also contains all potential cross-terms between the Sterimol parameters for both the substrate and ligand to examine potential synergistic relationships. A backward stepwise regression was performed, removing terms and optimizing the model based on f-tests of statistical significance for the model and p-tests for the individual coefficients.[6] The result of this regression is a simplified model shown as eq 2.

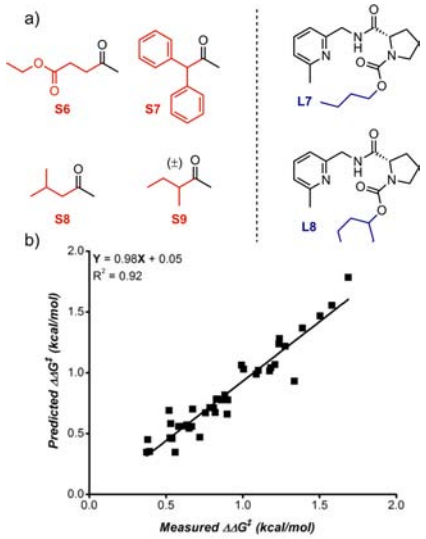
$$\Delta\Delta G^{\ddagger} = z0 + aB_{1L} + bB_{5L} + cL_L + dB_{1S} + fB_{5S} + gL_S + hB_{1L}B_{1S} + iB_{1L}B_{5S} + jB_{1L}L_S + kB_{5L}B_{1S} + mB_{5L}B_{5S} + nB_{5L}L_S + oL_LB_{1S} + pL_LB_{5S} + qL_LL_S \quad (1)$$

$$\Delta\Delta G^{\ddagger} = -2.19 + 1.47B_{1L} + 0.94B_{1S} - 0.27B_{1L}B_{1S} - 0.09B_{1L}B_{5S} \quad (2)$$

The dimensionality of eq 2 does not allow visualization by graphical means, but a plot of experimentally measured values for enantioselectivity, as well as those predicted by eq 2, is depicted in Figure 2c and demonstrates a high correlation between model and the experimentally observed enantioselectivity, with R^2 = 0.88. The model also passes the f-test at a 95% confidence level. The training set employed 30 unique ligand−substrate combinations, and the model is capable of describing the variation of these data using only five terms. The main characteristics of the model demonstrate the importance of the B_1 parameter on enantioselectivity. Large positive coefficients for B_1 are observed in the substrate and ligand dimensions. The two cross-terms are interesting. The $B_{1L}B_{1S}$ cross-term has a negative coefficient, indicating a small but negative effect on enantioselectivity. A possible explanation for the negative $B_{1L}B_{1S}$ steric effect is that the rate of reaction is decreased significantly when larger groups are present in both substrate and catalyst, which might magnify the effects of less selective background reactions eroding the observed enantioselectivity. The final cross-term is the $B_{1L}B_{5S}$ term, indicating that enantioselectivity is negatively affected, albeit modestly, by substrates with large B_5 values.

Validation. To evaluate the predictive power of the model described by eq 2, extensive independent validation was performed. Two new ligands, **L7** and **L8**, and four new substrates, **S6**−**S9**, were selected to validate eq 2 (Figure 3a). Each new ligand and substrate was evaluated in combination with all other substrates and ligands, respectively. The results are 41 measured enantioselectivities, which were compared with the values predicted by eq 2. The comparison between the predicted enantioselectivity and those observed experimentally for each new ligand−substrate combination is depicted in Figure 3b. A perfect model would possess a slope of 1.00, where deviation from unity represents decreasing predictive power for a model. The slope of 0.98 indicates that eq 2 is highly predictive. By comparison to the QSAR literature, similar models are considered predictive when slope values are

dx.doi.org/10.1021/ja4001807 | J. Am. Chem. Soc. 2013, 135, 2482−2485



c)

Substrate	Ligand	Measured ΔΔG‡	Predicted ΔΔG‡	Error
S6	L7	0.37	0.35	0.01
S6	L8	0.63	0.57	0.01
S7	L7	0.67	0.56	0.03
S7	L8	0.82	0.68	0.05
S8	L7	0.54	0.46	0.01
S8	L8	0.67	0.70	0.01
S9	L7	0.90	0.78	0.01
S9	L8	1.10	1.02	0.01

Figure 3. (a) Validation ligands and substrates. (b) Plot of experimental vs predicted enantioselectivity using the model defined in eq 2. (c) Table of new substrate and ligand combinations comparing measured vs predicted enantioselectivity.

between 0.6 and 1.4. The validation set includes eight pairings of a new ligand and a new substrate. Direct comparison of these completely interpolative values is shown in Figure 3c and highlights the efficacy of the model.

The validation substrate set includes racemic 3-methyl-2-pentanone, which contains a chiral center. Often in asymmetric catalysis, enantiomers of the substrate interact distinctively with a chiral catalyst, leading to increased rate for matched diastereomeric catalyst/substrate pairs and decreased rate for mismatched pairs. Subjecting **S9** to the ketone propargylation protocol led to equal generation of both diastereomers for all eight ligands evaluated. This general lack of kinetic resolution between catalyst and substrate alone is not compelling in a synthetic setting. However, chiral separation of the product diastereomers revealed that both were formed with the same enantiomeric ratio for each of the eight ligands evaluated (Figure 4). This suggests that the catalyst is not differentiating the substrate according to the adjacent carbons' configuration. Juxtaposed to the lack of diastereomeric resolution is the predictive power of eq 2, which predicts with reasonable accuracy the enantioselectivity for this substrate (Figure 4). The fact that eq 2 is heavily dependent on B_1 or proximal steric effects implies that the facial selectivity imparted by the catalyst is governed by this consequence independent of substrate chirality, although the steric difference between an ethyl and a methyl group is modest at best. Mechanistically, these results may suggest an open transition state, significantly reducing the number of hypothetical mechanisms of asymmetric induction.

Equation 2 was generated with a training set that included only methyl ketones, and the resultant model suggests that the substrate is orienting itself in such a manner to minimize steric interactions among its larger group and the carbamate substituent of the ligand. To expand the utility of the model,

Ligand	dr	Peak 1 er	Peak 2 er	Measured ΔΔG‡	Predicted ΔΔG‡	Error
L1	1:1	18:82	18:82	0.86	0.78	0.01
L2	1:1	19:81	20:80	0.83	0.78	0.01
L3	1:1	14:86	14:86	1.09	0.99	0.01
L4	1:1	16:84	16:84	1.00	1.03	0.01
L5	1:1	6:94	6:94	1.58	1.55	0.01
L6	1:1	5:95	5:95	1.69	1.78	0.01
L7	1:1	18:82	18:82	0.90	0.78	0.01
L8	1:1	14:86	14:86	1.10	1.02	0.01

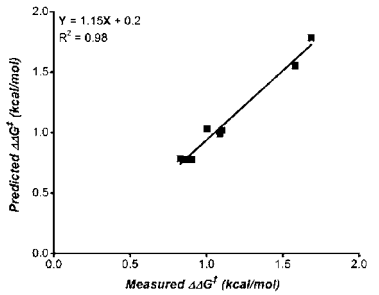


Figure 4. Evaluation of 3-methyl-2-pentanone (**S9**).

we evaluated two cyclic ketones, **S10** and **S11** (Figure 5). To predict the enantioselectivities for these cyclic ketones, the

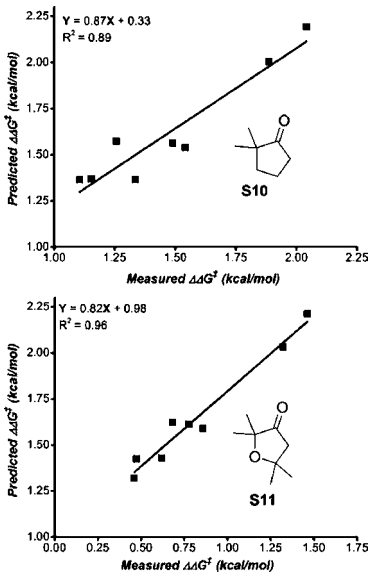


Figure 5. Evaluation of cyclic ketones.

existing model was revised to include the calculated Sterimol parameters for each side of the ketone, and the smaller substituent parameters were subtracted from the corresponding larger parameters. For the training set, this equates to the subtraction of the B_1, B_5, and L values for the methyl group. A new model was derived using the difference between Sterimol values of the substrate, and the only effect is a change to the coefficient values shown in eq 3.

$$\Delta\Delta G^{\ddagger} = -0.58 + 0.81B_{1L} + 0.94B_{1S} - 0.27B_{1L}B_{1S} - 0.087B_{1L}B_{5S} \quad (3)$$

The Sterimol values of the cyclic ketones could then be calculated and the difference taken and applied to eq 3. Figure 5 depicts the plots of predicted enantioselectivities from eq 3 for **S10** and the experimentally observed values. The results show that eq 3 is capable of predicting enantioselectivity with reasonable accuracy, validating the subtractive model. However, the predicted enantioselectivities for **S11** were much higher

dx.doi.org/10.1021/ja4001807 | J. Am. Chem. Soc. 2013, 135, 2482−2485

than observed (Figure 5). Evaluation of the predicted versus measured plot reveals good correlation of the predicted values but poor accuracy, suggesting that the source of this error is systematic, with each enantioselectivity being overestimated by ~1 kcal/mol. Because **S10** and **S11** present similar steric environments, especially in regard to the B_1 parameter, we explored a potential source of this systemic error. The purpose of this study was to evaluate the steric effects of ligand and catalyst simultaneously; to simplify the problem, we did not incorporate an electronic parameter in either our design matrix or our analysis, even though previous efforts suggest that electronic effects may be pertinent.[4g] To examine the potential electronic difference between these ketones, the carbonyl stretching frequencies were measured, and a significant disparity is observed (1733 cm^{-1} for **S10** and 1753 cm^{-1} for **S11**). The electronic difference between the ketones could be the source of this large systematic error, and ongoing studies are focused on the complex analysis associated with these combined effects.

In conclusion, through the application of Sterimol parameters and multivariate linear regression models, catalyst and substrate steric effects have been correlated. Extensive validation of these models resulted in excellent predictive power. Although this analysis was limited to steric interactions, the potential for applying a multivariate parametrization approach to include electronic effects could greatly enhance the application and breadth of the predictive power. Ultimately, a substrate scope could be designed using DoE principles to explore known reaction sensitivities, followed by regression analysis to generate models capable of predicting the performance of a wide range of substrates. A substrate scope defined by a model, rather than the simple or available substrates often reported, simplifies a key challenge in applying asymmetric catalysis: knowing how to extrapolate catalyst performance to substrate types not included in the original scope evaluation. Not only would this impact the synthetic user of the method, but the results would likely expose the key features in the origin of asymmetric induction. These are goals of our ongoing program.

■ ASSOCIATED CONTENT

Supporting Information

Experimental procedures, model development, and characterization data for new substances. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

sigman@chem.utah.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

K.C.H. thanks the ACS Division of Organic Chemistry and the University of Utah Graduate School for fellowships. This work was supported by the National Science Foundation (CHE-1110599).

■ REFERENCES

(1) (a) Kozlowski, M. C.; Waters, S. P.; Skudlarek, J. W.; Evans, C. A. *Org. Lett.* **2002**, *4*, 4391. (b) Kozlowski, M. C.; Panda, M. *J. Mol. Graphics Modell.* **2002**, *20*, 399. (c) Lipkowitz, K. B.; Kozlowski, M. C. *Synlett* **2003**, 1547. (d) Fristrup, P.; Tanner, D.; Norrby, P.-O. *Chirality* **2003**, *15*, 360. (e) Kozlowski, M. C.; Panda, M. *J. Org. Chem.* **2003**, *68*, 2061. (f) Jiang, C.; Li, Y.; Tian, Q.; You, T. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1876. (g) Melville, J. L.; Andrews, B. I.; Lygo, B.; Hirst, J. D. *Chem. Commun.* **2004**, 1410. (h) Melville, J. L.; Lovelock, K. R. J.; Wilson, C.; Allbutt, B.; Burke, E. K.; Lygo, B.; Hirst, J. D. *J. Chem. Inf. Model.* **2005**, *45*, 971. (i) Ianni, J. C.; Annamalai, V.; Phuan, P.-W.; Panda, M.; Kozlowski, M. C. *Angew. Chem., Int. Ed.* **2006**, *45*, 5502. (j) Chen, J.; Wen, J.; Li, M.; You, T. *J. Mol. Catal. A: Chem.* **2006**, *258*, 191. (k) Urbano-Cuadrado, M.; Carbo, J. J.; Maldonado, A. G.; Bo, C. *J. Chem. Inf. Model.* **2007**, *47*, 2228. (l) Jiang, C.; Li, D.; Wen, J.; You, T. *J. Mol. Model.* **2007**, *13*, 91. (m) Donoghue, P. J.; Helquist, P.; Norrby, P.-O.; Wiest, O. *J. Chem. Theory Comput.* **2008**, *4*, 1313. (n) O. Nilsson Lill, S.; Forbes, A.; Donoghue, P.; Verdolino, V.; Wiest, O.; Rydberg, P.; Norrby, P.-O. *Curr. Org. Chem.* **2010**, *14*, 1629. (o) Denmark, S. E.; Gould, N. D.; Wolf, L. M. *J. Org. Chem.* **2011**, *76*, 4337. (p) Denmark, S. E.; Gould, N. D.; Wolf, L. M. *J. Org. Chem.* **2011**, *76*, 4260. (q) Weill, N.; Corbeil, C. R.; De Schutter, J. W.; Moitessier, N. *J. Comput. Chem.* **2011**, *32*, 2878.

(2) (a) Tute, M. S. In *Comprehensive Medicinal Chemistry*; Hansch, C., Sammes, P. G., Taylor, J. B., Eds.; Pergamon Press: Oxford, 1990; Vol. *4*, p 1. (b) Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995. (c) Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 457. (d) Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 538. (e) Gonzalez, M. P.; Teran, C.; Saiz-Urra, L.; Teijeira, M. *Curr. Top. Med. Chem.* **2008**, *8*, 1606. (f) Verma, J.; Khedkar, V. M.; Coutinho, E. C. *Curr. Top. Med. Chem.* **2010**, *10*, 95.

(3) (a) Jacobsen, E. N.; Zhang, W.; Guler, M. L. *J. Am. Chem. Soc.* **1991**, *113*, 6703. (b) Lewis, C. A.; Gustafson, J. L.; Chiu, A.; Balsells, J.; Pollard, D.; Murry, J.; Reamer, R. A.; Hansen, K. B.; Miller, S. J. *J. Am. Chem. Soc.* **2008**, *130*, 16358. (c) Rodríguez-Escrich, S.; Reddy, K. S.; Jimeno, C.; Colet, G.; Rodríguez-Escrich, C.; Solà, L. s.; Vidal-Ferran, A.; Pericàs, M. A. *J. Org. Chem.* **2008**, *73*, 5340. (d) Zuend, S. J.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2009**, *131*, 15358. (e) Knowles, R. R.; Jacobsen, E. N. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 20678. (f) Knowles, R. R.; Lin, S.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2010**, *132*, 5030. (g) Gustafson, J. L.; Sigman, M. S.; Miller, S. J. *Org. Lett.* **2010**, *12*, 2794. (h) Uyeda, C.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2011**, *133*, 5062.

(4) (a) Jensen, K. H.; Sigman, M. S. *Angew. Chem., Int. Ed.* **2007**, *46*, 4748. (b) Miller, J. J.; Sigman, M. S. *Angew. Chem., Int. Ed.* **2008**, *47*, 771. (c) Sigman, M. S.; Miller, J. J. *J. Org. Chem.* **2009**, *74*, 7633. (d) Jensen, K. H.; Sigman, M. S. *J. Org. Chem.* **2010**, *75*, 7194. (e) Jensen, K. H.; Webb, J. D.; Sigman, M. S. *J. Am. Chem. Soc.* **2010**, *132*, 17471. (f) Harper, K. C.; Sigman, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2179. (g) Harper, K. C.; Sigman, M. S. *Science* **2011**, *333*, 1875. (h) Harper, K. C.; Bess, E. N.; Sigman, M. S. *Nat. Chem.* **2012**, *4*, 366.

(5) (a) Verloop, A. In *Drug Design*; Ariens, E. J., Ed.; Academic Press: New York, 1976; Vol. *III*, p 133. (b) Verloop, A., Tipker, J. In *Biological Activity and Chemical Structure*; Buisman, J. A., Ed.; Elsevier: Amsterdam, 1977; p 63. (c) Verloop, A., Tipker, J. In *QSAR in Drug Design and Toxicology*; Hadzi, D., Jerman-Blazic, B., Eds.; Elsevier: Amsterdam, 1987; p 97. (d) Verloop, A. In *IUPAC Pesticide Chemistry*; Miyamoto, J., Ed.; Pergamon: Oxford, 1983; Vol. *1*, p 339.

(6) Deming, S. N.; Morgan, S. L. *Experimental Design: A Chemometric Approach*, 2nd revised and expanded edition.; Elsevier: Amsterdam, 1993.