Published in final edited form as:

Angew Chem Int Ed Engl. 2017 November 06; 56(45): 14080–14084. doi:10.1002/anie.201707644.

Multidimensional Correlations in Asymmetric Catalysis via Parameterization of Uncatalyzed Transition States

Dr. Manuel Orlandi^[a], Prof. F. Dean Toste^[b], and Prof. Matthew S. Sigman^[a]

^[a]Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, USA

^[b]Department of Chemistry, University of California, Berkeley, California 94720, United States

Abstract

The study of the chiral anion phase transfer (CAPT) catalyzed oxidative amination of tetrahydroisoquinolines via multidimensional correlation analysis (MCA) is revisited. The parameterization of the transition states (TSs) for the uncatalyzed reaction, the introduction of conformational descriptors, and the use of computed interaction energies and distances as parameters allowed access to a considerably simplified mathematical correlation of substrate and catalyst structure to enantioselectivity. The equation obtained is suggestive of key interactions occurring at the TS. Specifically, the CAPT catalyst is proposed to coordinate the intermediate iminium cation by $P=O \cdots H-O$ hydrogen bond and by $N \cdots H-C$ electrostatic interaction. The conformational freedom of the substrate's benzyl substituent is also found to be important in order to provide an efficient mode of molecular recognition.

Graphical Abstract

The study of the chiral anion phase transfer catalyzed oxidative amination of tetrahydroisoquinolines via multidimensional correlation analysis is revisited. The parameterization of the transition states for the uncatalyzed reaction, the introduction of conformational descriptors, and the use of computed interaction energies and distances as parameters allowed access to a considerably simplified mathematical correlation of substrate and catalyst structure to enantioselectivity.

Keywords

multidimensional correlation analysis; molecular parametrization; free-energy relationships; chiral anion phase transfer catalysis; asymmetric catalysis

Correspondence to: Matthew S. Sigman.

Supporting information for this article is given via a link at the end of the document.

Multidimensional correlation analysis (MCA) has emerged as a mechanistic tool for transition state (TS) interrogation in asymmetric catalysis.^[1] Using this technique, a polynomial equation that correlates the reaction outcome with the chemical descriptors of a catalyst and/or substrate set can be derived. The resultant mathematical model can be used to extrapolate better catalysts through virtual screening thus improving the reaction performance.^[2] Additionally, analysis of the parameters included in the equation provides information about the reaction mechanism.^[2b, 3] Historically, limited parameters were available to probe such mechanistic considerations including Hammett σ ,^[4] Tolman cone angle,^[5] Charton^[6] and Taft parameters,^[7] yet in recent years modern physical organic descriptors have been reported,^[8] which have ultimately led to the rationalization of reactions of enhanced complexity.^[3b, 9] Even with these advances and the ability to perform multi-variate analysis of intricate processes for predictive purposes, there have been a number of cases wherein the sophistication of the reaction does not lend itself to straightforward analysis. As an example, we reported the correlation of both substrate and catalyst descriptors to enantioselectivity in the chiral anion phase transfer (CAPT)^[10] catalyzed oxidative amination of tetrahydroisoquinolines depicted in Figure 1A. In this report, a data-intensive approach was used in order to obtain a complex mathematical model that allowed extrapolation of the improved catalyst **1a** (Figure 1B and 1C). In this model, several IR^[9b] intensity and frequency parameters for the triazole moiety suggested the presence of noncovalent interactions (NCIs) between the ring and the substrate in the TS (Figure 1C), yet it was not possible to build a descriptive picture of the interactions at the heart of asymmetric catalysis. Thus, we envisioned applying a unique strategy to develop a parameter set suited to this complex problem. Specifically, we envisioned applying our recently introduced computed interaction energies $E\pi$ and distances $D\pi$,^[3d] which could provide specific insight into the influential NCIs. Of greater importance, we hypothesized that the readily computed transition state of the corresponding uncatalyzed cyclization would represent an optimal platform for parameters acquisition due to its resemblance to the catalyzed reaction TS. We posited that this effort would deepen our mechanistic understanding regarding the stereodefining step and the specific molecular recognition events required to obtain high enantioselectivity.^[2b, 11] Herein, we describe the application of this approach to significantly simplify the equation describing enantioselectivity and provide a detailed overview of how asymmetric catalysis is effectively achieved.

Due to the influence that the catalyst's triazole and the substrate's benzyl group have on the reaction stereochemical outcome,^[2b] the generic interaction in Figure 1C was previously hypothesized. However, other portions of cation **3** may be recognition elements for the catalyst. Hence, interaction energies and distances^[3d] were computed at the B97D/def2TZVP level of theory^[12] using three different probes: benzene **PhH** (E/D_{PhH}), iminium cation **ImC** (E/D_{ImC}), and anilinium cation **AnC** (E/D_{AnC} , Figure 2A). These probes describe the possible interaction of the triazole with three different moieties of intermediate iminium **3** (Figure 2A). Specifically, E/D_{PhH} , E/D_{ImC} and E/D_{AnC} , respectively, account for the possible interactions with the neutral benzyl group, with the alkyl portion next to the charged N-atom, and with the two electron poor aryl groups. Other parameters for catalysts **1a–1m** were also acquired according to a DFT molecular model that we recently introduced for chiral phosphate catalysts (Figure 2C).^[3d, 10i] IR frequencies and intensities (ν/i), the

dihedral angle α , and Sterimol parameters (**B1**, **B5** and **L**) for this molecular model were calculated at the M06-2X/def2TZVP level of theory.^[13]

The dependence of the reaction *ee* on the CAPT catalyst was investigated by correlation of the selectivities (expressed as $G^\ddagger$ in kcal/mol) measured for the cyclization of substrate **2a** (Figure 2B). Among all the parameters collected, **sen**(α) (sine of the dihedral angle α in the catalyst, Figure 2C) was the only single parameter that provided a qualitative trend ($R^2=0.66$, see SI), suggesting the importance of geometric and/or steric requirements to access high *ee*. When a multidimensional correlation was determined, the mathematical model in Figure 2C was obtained ($R^2=0.87$, L1O=0.74, L1O = Leave-1-Out cross validation). This equation is composed of three terms: **sen**(α), which accounts for geometrical/steric requirements; **vpos_y**, which describes the phosphate's coordination ability (or Lewis basicity); and the cross term **E_{ImC}·D_{ImC}**, which suggests the presence of a NCI between the triazole ring and the substrate alkyl portion.

Hence, based on the inclusion of these parameters in the multivariate model (Figure 2C), we hypothesized that the catalyst coordinate intermediate **3** through a dual binding mode. Specifically, the phosphate and the triazole ring could engage the amide group and the electron poor alkyl chain by H-bond and electrostatic interactions (**vpos_y** and **E_{ImC}·D_{ImC}**). The presence of bulky groups in a specific region of space (**sen**(α)) would then ensure discrimination of one of the enantiotopic faces in the TS. Additional evidence for the hypothesized coordination mode arises from the computational study of the uncatalyzed reaction. Transition states **TS-A** and **TS-B** were calculated at the M06-2X/def2TZVP level of theory. They differ in the tautomeric form in which the amide group acts as a nucleophile. In **TS-A**, the amide is in iminolic form and presents a free OH group available for coordination. In **TS-B**, the amide reacts in its most stable tautomeric form, yet shows distortion of the functional group out of planarity (Figure 3A). Thus, **TS-A** is favored by 5.9 kcal/mol, which suggests that intermediate **3** likely reacts in its iminol form (Figure 3A).

Evaluation of the electron density map for **TS-A** highlights the OH proton and the alkyl chain next to the iminium cation as the most electron-poor regions of the structure (Figure 3B). Thus, in order for these regions to participate in NCIs, it is likely that they will be matched with electron-rich regions of the catalyst. Calculation of the optimized geometry and electron density map for catalyst **1b** emphasized two important points: i) the chiral pocket of the catalyst is not accessible due to steric hindrance, and ii) the most electron rich regions of the catalyst are the phosphate and the triazole groups (Figure 3B). Hence, on the basis of this simple computational analysis and of the equation in Figure 2C, we posited the presence of two main interactions that govern the coordination of the reactive intermediate by the catalyst in the transition state. The P=O...H-O hydrogen bond provides coordination and activation of the nucleophilic amide as the catalyst acts as a Brønsted base (Figure 3A), and the electrostatic interaction between the triazole and the alkyl portion ensure optimal orientation for high stereoselectivity (Figure 3C).^[14] Indeed, when such a dual coordination mode is not allowed due to the substitution of the triazole with a pyrazole or imidazole ring, the selectivity decreases (compare **1m** and **1l,1k**, Figure 2B). In other words, the catalyst interacts with intermediate **3** through bidentate binding. Consistent with this hypothesis, the

distance between the two electron poor regions in the substrate and the two electron rich regions in the catalyst are relatively matched (5.0–6.6 Å and 4.9–6.8 Å, Figure 3D).

After assessing the role of the catalyst in this transformation, we turned our attention to the dependence of the reaction *ee* on the substrate. Although the reaction was found to proceed mainly under catalyst control, the perturbation of the substrate's benzyl group (especially in the 2,6-positions) resulted in a $G^\ddagger$ variation of up to 1.5 kcal/mol. Since phosphate **1** presumably coordinates iminium cation **3** via NCIs, the reagent geometry in the TS of the catalyzed and the uncatalyzed reaction is similar, and the stereoselectivity is due to enhanced structural matching of the catalyst with one of the two enantiomers of such TSs. Thus, the TS of the uncatalyzed reaction represents an optimal structure for substrate parametrization. Additionally, it has fewer conformers than the relative reagent structure as the presence of the forming C–N bond reduces the molecular degrees of freedom. Hence, descriptors for substrates **2a–2k** (Figure 4A) were calculated from the uncatalyzed TSs at the M06-2X/def2TZVP level of theory. The computed parameters include IR frequencies and intensities (ν/i , including the imaginary frequency associated to the C–N bond formation), Sterimol parameters (**B1**, **B5** and **L**), and NBO charges.

As molecular recognition depends on the substrate geometry, the selectivity is due to the ability of the catalyst to match with one of the two enantiomeric forms of the uncatalyzed TS in a specific conformation. Hence, we reasoned that the conformational freedom of the benzyl group could be important for optimal interactions. Thus, the energy difference E_{AB} between the two main conformers of the computed uncatalyzed TSs (structures **A** and **B**, Figure 4B) was included as a parameter in addition to IR, Sterimol and NBO charge descriptors. Interestingly, E_{AB} provided good single-parameter correlations (Figure 4B) with $G^\ddagger$ for both catalyst **1b** and **1j**, which present different steric environments. However, while such correlation for substrates **2e–2k** shows the same trend for different catalysts, substrates **2a–2d** are outliers for bulky catalysts **1a–1c** (Figure 4B).

With the new set of parameters determined, a multidimensional model from the combination of catalysts **1a–1j** and substrates **2a–2k** was identified (see SI for the full list). This model includes 103 data points (70 for the training set and 33 for the external validation set) and includes only six terms (Figure 4C). A 13-term equation (Figure 1C) was previously required for the description of this complex reaction system.^[2b] The parameters included in the equation are as follows: $\text{sen}(\alpha)$, ν_{POSy} and E_{ImC} for the catalyst, which also appeared in the equation in Figure 2B; E_{AB} and $E_{AB} \cdot \text{sen}(\alpha)$, which account for the conformational freedom of the substrate (E_{AB}) and for the ability of the catalyst to recognize the substrate conformation based on its geometry ($E_{AB} \cdot \text{sen}(\alpha)$); ν_{CN} (stretching frequency of the C=N bond in the uncatalyzed TS), which may describe the substrate nucleophilicity. The multivariate model presents $R^2=0.83$ and both external and cross validations (LKO, Leave K Out tests) assess its robustness. Hence, the parameters involved in the full model suggest the interaction mode depicted in Figure 3C. Since the reaction proceeds mainly under catalyst control,^[2b] parameters for the catalysts have larger coefficients. Specifically, the catalyst's ability to coordinate iminium **3** via H-bond (ν_{POSy}) and N...H–C electrostatic interaction (E_{ImC}), and the steric/geometric properties of the catalyst's substituents ($\text{sen}(\alpha)$) play the most important role in the success of the reaction. Additionally, the presence of E_{AB}

suggests a specific molecular recognition, in which the catalyst typically prefers to react with substrates in conformation **A**.

In conclusion, the mechanism of the CAPT catalyzed oxidative amination of tetrahydroisoquinolines has been revisited. The mathematical model previously developed by our groups provided improvement of the reaction performance by extrapolation of a better performing catalyst (**1a**). In this previous study, mechanistic suggestions about the role of the triazole group could be obtained, yet the complex equation (13 terms with 9 parameters) did not allow a clear understanding of the mode of coordination between the catalyst and the reactive intermediate **3**.^[2b] The strategy presented herein applies computed interaction energies and corresponding distances to access more detailed mechanistic information. The use of the uncatalyzed reaction TSs has been introduced as a new general strategy to compute parameters for intramolecular processes. Similarly, conformational relative energies have also been introduced in order to describe the conformational freedom of substrate **2**. This new set of parameters allowed us to obtain a significantly simplified multidimensional model, which clearly suggests the role of the catalyst's triazole ring in the TS. In particular, it was hypothesized to be involved in the coordination of the substrate via electrostatic interaction with the alkyl portion next to the charged N-atom of intermediate **3**. Additionally, the conformational freedom of the substrate was proposed to be important for optimal molecular recognition from the catalyst. We foresee the new parameters and parametrization strategy reported here to become a useful tool in MCA, as they allow access to detailed mechanistic information even for complex catalytic systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the NIH (1 R01 GM121383) for support of this work. M.O. thanks the Ermenegildo Zegna Group for a postdoctoral fellowship. Computations were conducted at the Center for High Performance Computing (CHPC) of the University of Utah. We thank Dr. M. J. Hilton for helpful discussions.

References

1. a) Sigman MS, Harper KC, Bess EN, Milo A. *Acc Chem Res.* 2016; 49:1292. [PubMed: 27220055] b) Yang C, Zhang E-G, Li X, Cheng J-P. *Angew Chem Int Ed.* 2016; 55:6506. *Angew Chem.* 2016; 128:6616. c) Piou T, Romanov-Michailidis F, Romanova-Michaelides M, Jackson KE, Semakul N, Taggart TD, Newell BS, Rithner CD, Paton RS, Rovis T. *J Am Chem Soc.* 2017; 139:1296. [PubMed: 28060499] d) Yang C, Wang J, Liu Y, Ni X, Li X, Cheng JP. *Chem Eur J.* 2017; 23:5488. [PubMed: 28244165] e) Huang H, Zong H, Bian G, Yue H, Song L. *J Org Chem.* 2014; 79:9455. [PubMed: 25222868]
2. a) Harper KC, Sigman MS. *Science.* 2011; 333:1875. [PubMed: 21960632] b) Milo A, Neel AJ, Toste FD, Sigman MS. *Science.* 2015; 347:737. [PubMed: 25678656] c) Zhang C, Santiago CB, Crawford JM, Sigman MS. *J Am Chem Soc.* 2015; 137:15668. [PubMed: 26624236]
3. a) Chen ZM, Hilton MJ, Sigman MS. *J Am Chem Soc.* 2016; 138:11461. [PubMed: 27571167] b) Niemeyer ZL, Milo A, Hickey DP, Sigman MS. *Nat Chem.* 2016; 8:610. [PubMed: 27219707] c) Guo JY, Minko Y, Santiago CB, Sigman MS. *ACS Catal.* 2017:4144. d) Orlandi M, Coelho JAS, Hilton MJ, Toste FD, Sigman MS. *J Am Chem Soc.* 2017; 139:6803. [PubMed: 28475315]
4. Hammett LP. *J Am Chem Soc.* 1937; 59:96.

5. Tolman CA. Chem Rev. 1977; 77:313.
6. Charton M. J Am Chem Soc. 1975; 97:1552.
7. Taft RW. J Am Chem Soc. 1953; 75:4538.
8. a) Clavier H, Nolan SP. Chem Commun. 2010; 46:841. b) Dröge T, Glorius F. Angew Chem Int Ed. 2010; 49:6940. Angew Chem. 2010; 122:7094. c) Wu K, Doyle AG. Nat Chem. 2017; doi: 10.1038/nchem.2741
9. a) Harper KC, Bess EN, Sigman MS. Nat Chem. 2012; 4:366. [PubMed: 22522256] b) Milo A, Bess EN, Sigman MS. Nature. 2014; 507:210. [PubMed: 24622199] c) Mougél V, Santiago CB, Zhizhko PA, Bess EN, Varga J, Frater G, Sigman MS, Copéret C. J Am Chem Soc. 2015; 137:6699. [PubMed: 25938259] d) Santiago CB, Milo A, Sigman MS. J Am Chem Soc. 2016; 138:13424.
10. a) Rauniyar V, Lackner AD, Hamilton GL, Toste FD. Science. 2011; 334:1681. [PubMed: 22194571] b) Phipps RJ, Hiramatsu K, Toste FD. J Am Chem Soc. 2012; 134:8376. [PubMed: 22574822] c) Lackner AD, Samant AV, Toste FD. J Am Chem Soc. 2013; 135:14090. [PubMed: 24025122] d) Phipps RJ, Toste FD. J Am Chem Soc. 2013; 135:1268. [PubMed: 23330962] e) Nelson HM, Reisberg SH, Shunatona HP, Patel JS, Toste FD. Angew Chem Int Ed. 2014; 53:5600. Angew Chem. 2014; 126:5706. f) Yang X, Phipps RJ, Toste FD. J Am Chem Soc. 2014; 136:5225. [PubMed: 24684209] g) Zi W, Wang YM, Toste FD. J Am Chem Soc. 2014; 136:12864. [PubMed: 25203796] h) Nelson HM, Williams BD, Miró J, Toste FD. J Am Chem Soc. 2015; 137:3213. [PubMed: 25723255] i) Yamamoto E, Hilton MJ, Orlandi M, Saini V, Toste FD, Sigman MS. J Am Chem Soc. 2016; 138:15877. [PubMed: 27960315] j) Yang Z, He Y, Toste FD. J Am Chem Soc. 2016; 138:9775. [PubMed: 27472195]
11. Neel AJ, Hehn JP, Tripet PF, Toste FD. J Am Chem Soc. 2013; 135:14044. [PubMed: 24025163]
12. Grimme S. J Comput Chem. 2006; 27:1787. [PubMed: 16955487]
13. a) Valero R, Gomes JRB, Truhlar DG, Illas F. J Chem Phys. 2008; 129:124710. [PubMed: 19045051] b) Zhao Y, Truhlar DG. Theor Chem Acc. 2008; 120:215.
14. a) Zimmer LE, Sparr C, Gilmour R. Angew Chem Int Ed. 2011; 50:11860. Angew Chem. 2011; 123:12062. b) Seguin TJ, Wheeler SE. ACS Catal. 2016; 7:222. c) Seguin TJ, Wheeler SE. Angew Chem Int Ed. 2016; 55:15889. Angew Chem. 2016; 128:16121. d) Seguin TJ, Wheeler SE. ACS Catal. 2016; 6:2681. e) Wheeler SE, Seguin TJ, Guan Y, Doney AC. Acc Chem Res. 2016; 49:1061. [PubMed: 27110641] f) Duarte F, Paton RS. J Am Chem Soc. 2017; 139:8886. [PubMed: 28582626]

Figure 1.

A) Case study: oxidative amination of tetrahydroisoquinolines. B) CAPT catalysts considered in the present study. C) Multivariate model previously reported and interaction hypothesized between the substrate benzyl group and the catalyst's triazole ring.

Figure 2.

A) Interaction energies and distances calculated in the present study (E/D_{PhH} , E/D_{ImC} and E/D_{AnC}). B) Selectivities obtained in the oxidative cyclization of substrate **2a** with CAPT catalysts **1a–1m**. C) Parameters calculated for catalysts **1a–1m** and multivariate model obtained.

Figure 3.

A) Relative energy of the tautomeric **TS-A** and **TS-B**. B) Electron density maps for **TS-A** and **1b**. C) Hypothesized TS for the reaction. D) Distance between electronpoor regions in **TS-A** and electronrich regions in catalyst **1b**.

Figure 4.

A) Substrates **2a–2k** included in the present study. B) Conformational freedom of substrate **2a**. C) Full multidimensional model obtained from 103 data points, and parameters included in the model.