

# Using Stochastic Spectral Moments to Predict In-Out Perturbations in the Synthetic Pathways involving Asymmetric Catalysis of Intra-molecular Carbolithiations

Humberto González-Díaz<sup>1,2,\*</sup>; Sonia Arrasate<sup>1</sup>, Asier Gómez-SanJuan<sup>1</sup>, Nuria Sotomayor<sup>1</sup>, Esther Lete<sup>1</sup>, Alejandro Speck-Planche<sup>3,4</sup>, and M. N. D. S. Cordeiro<sup>4,\*</sup>

<sup>1</sup> Department of Organic Chemistry II, University of the Basque Country UPV/EHU, 48940, Leioa, Spain

<sup>2</sup> IKERBASQUE, Basque Foundation for Science, 48011, Bilbao, Spain

<sup>3</sup> Department of Applied Physics, USC, 15782, Santiago de Compostela, Spain.

<sup>4</sup> REQUIMTE/Department of Chemistry and Biochemistry, University of Porto, 4169-007 Porto, Portugal.

**Abstract.** We develop a new QSPR-Perturbation theory model for Synthetic Pathways involving Asymmetric Catalysis of Intra-molecular Carbolithiations. In doing so, we used calculated spectral moments of markov electron delocalization matrix for all molecules involved in the reactions. These descriptors and physical variables (temperature etc.) were used to calculate perturbation operators that account for perturbations in the synthetic reactions. The model classifies correctly >100,000 pairs of intra-molecular carbolithiations with high Accuracy (Ac), Sensitivity (Sn), and Specificity (Sp). The model predicts probabilities of variations in the yield and enantiomeric excess of reactions due to at least one perturbation in boundary conditions (solvent, temperature, temperature of addition, or time of reaction). The model is useful to predict the effect over synthetic routes of changes in chemical structure (connectivity structure and/or chirality patterns in substrate, product, electrophile agent, organolithium, and ligand of the asymmetric catalyst).

**Keywords:** asymmetric catalysis; carbolithiations; perturbation theory; QSPR; spectral moments

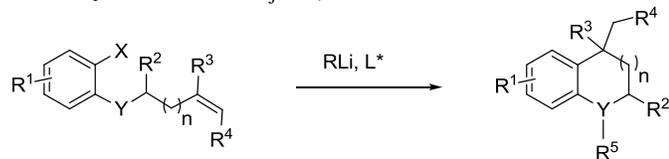
\***Address correspondence to.** González-Díaz H: Department of Organic Chemistry II, University of the Basque Country (UPV/EHU), 48940, Bilbao, Spain. Email: [humberto.gonzalezdiaz@ehu.es](mailto:humberto.gonzalezdiaz@ehu.es)

## Introduction

A particular class of reaction of high interest in organic synthesis is the so called carbolithiation. The carbolithiation reaction offers an attractive pathway for the efficient construction of new carbon-carbon bonds by addition of an organolithium reagent to non-activated alkenes or alkynes, with the possibility of introducing further fictionalization on the molecule [1-7]. When alkenes are used, up to two contiguous stereogenic centers may be generated, which may be controlled by using chiral ligands for lithium, and so opening new opportunities for application of this methodology to asymmetric synthesis. The naturally occurring alkaloid (-)-sparteine is the most widely used chiral ligand in enantioselective carbolithiation reactions. These reactions can be carried out either in an inter- or intra-molecular fashion, though only a few examples have been described, due to the difficulty of enantiofacial differentiation for a non-activated alkene. The intra-molecular variant of this reaction has been applied mainly with alkyl- and alkenyllithium intermediates, though there are also examples of a cycloisomerization reaction for alkenyl-substituted aryllithium intermediates,

generated by a metal-halogen exchange [1-6]. In particular, intra-molecular carbolithiation is well suited for the diastereoselective construction of five-membered rings through a *5-exo-trig* cyclization process, as has been shown in the synthesis of both carbocycles [8, 9] and heterocycles [10-12]. However, there are few precedents for the formation of six-membered cycles by the carbocyclization of unactivated double bonds, and its application to medium-sized rings has not been reported yet. To achieve a *6-exo* cyclization it is necessary to use a substituted alkene to stabilize the resulting organolithium compound. In this work we are going to develop a QSPR-perturbation model for a very large set of input-output perturbations in the reactions described in **Figure 1**.

Very recently Gonzalez-Díaz *et al.* [13] formulated a general-purpose perturbation theory or model for multiple-boundary QSPR/QSAR problems. The general picture considered as starting point for this theory is the following. Let be a set of molecules  $m_{qi}$  that undergoes a transformation process  $tp$  experimentally measured under a set of boundary conditions  $c_j \equiv (c_0, c_1, c_2, c_3 \dots c_n)$ .



$R^1 = \text{CH}_3, \text{F}, \text{OBn}, \text{OCH}_3$

$R^2 = \text{H}, \text{Ph}$

$R^3 = \text{H}, \text{Pr}, \text{Ph}, \text{CH}_2\text{OH}, \text{CH}_2\text{OMe}, \text{CH}_2\text{OTIPS}, \text{CH}_2\text{SPh}, \text{CH}_2\text{N}(\text{Me})_2$

$R^4 = \text{H}, \text{CON}(\text{Et})_2, \text{CONMe}(\text{OMe})$

$R^5 = \text{CH}_3, \text{Bn}, \text{allyl}$

$X = \text{Br}, \text{I}$

$Y = \text{N}, \text{O}$

$n = 0, 1$

**Figure 1.** General scheme for intramolecular carbolithiations studied in the present work

The  $tp$  involve at least chemical reaction ( $cr$ ) and/or one physicochemical process ( $pp$ ) like organic synthesis reactions, biochemical reactions, or drug metabolism and/or physicochemical processes like drug-target interaction, self-aggregation in complex systems, or absorption, distribution, and excretion in a living organism. It means that the  $tp$  involve at least one  $cr$  that change the structure and/or at least one  $pp$  that shift the state of this set of  $i^{\text{th}}$  molecules ( $m_{qi}$ ). We also have to consider that a  $tp$  may involve more than one input/output molecules; which play different distinguishable roles or functions  $f^{\text{th}}$  in the  $cr^{\text{th}}$  reaction or  $pp^{\text{th}}$  process. Another important fact is the existence of multiple classes  $s^{\text{th}}$  of output efficiency parameters  $\varepsilon_s(m_{qi})$  to quantify the efficiency with which the molecules are transformed by the  $cr^{\text{th}}$  reaction or  $pp^{\text{th}}$  process involved in the transformation process  $tp^{\text{th}}$ . Our main aim is to predict the shift or change in a function of the output efficiency  $\Delta\lambda(\varepsilon_{ij}) = \lambda(\varepsilon_{ij})_{\text{new}} - \lambda(\varepsilon_{ij})_{\text{ref}}$  that takes place after a change, variation, or perturbation ( $\Delta V$ ) in the structure and/or boundary conditions of a peptide of reference.

## Results

In this particular case we constructed a two sets of state information functions  $\lambda(c_j)$  and  $\lambda_5(i m_q)_g$ . The functions of type  $\lambda(c_j)$  quantify information about boundary conditions that apply to all the reactions system. The two functions of type  $\lambda(c_j)$  used here

were the time of reaction  $t_{\text{reac}}$  and the temperature of reaction  $T_{\text{reac}}$ . On the other hand, the functions of type  $\lambda_5(i m_q)_g$  quantify structural, physical, and/or operational boundary conditions for one specific class of molecular entities  $m_q$ . This second set of functions has the general form  $\lambda_5(i m_q)_g = \pi_5(i m_q)_g \cdot f(c_j)$ . In this equation  $m_1 = \text{substrate (S)}$ ,  $m_2 = \text{organolithium reagent (RLi)}$ ,  $m_3 = \text{ligand of asymmetric catalyst (L)}$ ,  $m_4 = \text{solvent (V)}$ ,  $m_5 = \text{electrophiles (E)}$ , and  $m_6 = \text{product (P)}$ ; are the molecular entities that play different roles in the reaction. The symbol  $\pi_5$  stands for the spectral moment of order five of the stochastic matrix of electron delocalization ( ${}^1\Pi$ ), weighted by the electronegativity. The parameter quantifies both structural information in terms of bond connectivity and electronic distribution. Electronic information is measured as the probability of self-return of electrons to the original atom after interactions with to to its closer neighbour atoms placed at topological distance  $d \leq 5$ . This parameter can be calculated very fast for large libraries of compounds using the software MARCH-INSIDE. In recent works we have reviewed and discussed the theoretical basis and applications of MARCH-INSIDE in organic and medicinal chemistry [14, 15]. The subscript  $g$  accounts for the presence of groups of atoms including  $g_1 = \text{total atoms}$ ,  $g_2 = \text{non-saturated carbons}$ , or  $g_3 = \text{heteroatoms}$ . The symbol  $f(c_j)$  is a function that depends on some additional structural and/or physical boundary conditions. The specific state information functions for the different molecular entities are the following. For substrate:  $\lambda_5(\text{Subs})_m = \pi_5(\text{Subs})_m \cdot e^{[p(\text{SC})]}$ , being  $p(\text{SC})$  the probability associated with the chirality of the substrate. For RLi:  $\lambda_5(\text{RLi})_g = \pi_5(\text{RLi})_g \cdot T_{\text{add}} \cdot p_{\text{add}} \cdot e^{c(\text{RLi})}$ , where  $T_{\text{add}}$  is the temperature of addition of the organolithium,  $p_{\text{add}}$  is the priori probability associated with the method of addition of the organolithium derivative. The symbol  $c(\text{RLi})$  is the equivalent concentration of the organolithium reagent. For ligands the function is:  $\lambda_5(\text{L})_m = \pi_5(\text{L})_g \cdot p(\text{LC}) \cdot e^{c(\text{L})}$ , being  $p(\text{LC})$  the probability associated with the chirality of the ligand. The symbol  $c(\text{Lig})$  is the equivalent concentration of the ligand. For solvent:  $\lambda_5(\text{S})_g = \pi_5(\text{S})_g$ . For electrophiles:  $\lambda_5(\text{E})_g = \pi_5(\text{E})_g$ . For product:  $\lambda_5(\text{P})_g = \pi_5(\text{P})_g \cdot e^{[p(\text{PC}) \cdot \delta]}$ , being  $p(\text{PC})$  the probability associated with the chirality of the product. The symbol  $\delta$  is an identifier of the type of measure of efficiency  $\delta \varepsilon_i(c_j)$  of reaction used;  ${}^1\varepsilon_i(c_j)$

= yield (%) and  $^{-1}\varepsilon_i(c_j)$  = enantiomeric excess  $ee(\%)$ . This variable takes value of 1 when the yield is considered and -1 when enantiomeric excess is determined. We found a model useful to predict the efficiency of the new reaction  $\delta\varepsilon_i(c_j)_{nr}$  given the efficiency of the reference reaction  $\delta\varepsilon_i(c_j)_{rr}$  and the values of the perturbation terms  $\Delta\lambda_5^i(m_q) = \lambda_5^i(m_q)_{nr} - \lambda_5^i(m_q)_{rr}$  for the state information functions. The equation of this model is the discrimination function:

$$\lambda(\delta\varepsilon_i(c_j))_{nr} = 3.000 \cdot \lambda(\delta\varepsilon_i(c_j))_{rr} + 7.000 \cdot \Delta T_{\text{reac}} - 13.000 \cdot \Delta t_{\text{reac}} + 16.000 \cdot \Delta\lambda_5(S)_{\text{Het}} + 8.000 \cdot \Delta\lambda_5(\text{RLi})_{\text{Tot}} - 9.000 \cdot \Delta\lambda_5(L)_{\text{Cin}} + 0.521 \cdot \Delta\lambda_5(V)_{\text{Cin}} + 4.000 \cdot \Delta\lambda_5(E)_{\text{Het}} + 2.000 \cdot \Delta\lambda_5(P)_{\text{Het}} - 1.00$$

N = 86894  $\chi^2 = 38237.11$  p-level < 0.001

A very notable feature of this model is its ability to predict very well a high number of perturbations in a large library of intramolecular carbolithiations (See **Table 1**).

Data set	Training		Prediction	
	Negative	Positive	Negative	Positive
Total	82817	4077	27383	1414
Correct	80911	3982	26761	1393
Wrong	1906	95	622	21
Correct (%)	97.70	97.67	97.73	98.51
Wrong (%)	2.30	2.33	2.27	1.49
Accuracy	97.70		97.77	

In addition, the model is able to predict more than one output property. Note that, if we make the substitution  $^1\varepsilon_i(c_j)_{rr} = \text{yld}(c_j)_{rr}$  in the model we can predict the probability with which the new reaction has a yield  $\text{yld}(c_j)_{nr} \geq 95\%$ . On the other hand, if we substitute  $^{-1}\varepsilon_i(c_j)_{rr} = ee_i(c_j)_{rr}$  in the model we can predict the probability with which the new reaction

produce an enantiomeric excess  $ee_i(c_j)_{nr} > 90\%$  of the  $i^{\text{th}}$  product. This type of QSRR-Perturbation model may predict the formation of different chirality patterns in the product of the new reaction. This prediction depends on the chirality patterns of the ligand of the asymmetric catalyst of this reaction. The prediction depends also on the chirality patterns of both ligand and products of the reaction of reference.

## Conclusions

It is possible to use the spectral moments indices calculated with MARCH-INSIDE as inputs of statistical analysis in order to develop general QSPR models that predict the results (in terms of yield and enantiomeric excess) for multiple in-out perturbations in intramolecular carbolithiations. These models may include perturbations in a very high number of input-output variables like (time, temperature, solvent, catalyst, assay, pharmacological experimental measures, molecular and cellular targets, and many others).

## Acknowledgments

We wish to thank the Ministerio de Ciencia e Innovación (CTQ2009-07733) and Universidad del País Vasco/Euskal Herriko Unibertsitatea (UFI11/22, GIU 0946) for their financial support. Technical and personnel support provided by Servicios Generales de Investigación SGIker (UPV/EHU, MICINN, GV/EJ, ERDF and ESF) is gratefully acknowledged. We also thank UPV/EHU for a postdoctoral grant (A.G). A. Speck-Planche acknowledges the Portuguese Fundação para a Ciência e Tecnologia (FCT) and the European Social Found for financial support (SFRH/BD/77690/2011).

## References

- [1] Marek I. Enantioselective carbometallation of unactivated olefins. 1999.
- [2] Bailey WF, Fair JD. Intramolecular carbolithiation cascades as a route to a highly strained carbocyclic framework: competition between 5-exo-trig ring closure and proton transfer. *Beilstein J Org Chem*, 2013; 9: 537-43.
- [3] Rappoport Z, Marek I, Fañanás FJ, Sanz R. Intramolecular Carbolithiation Reactions. 295-379.
- [4] A.M. H, O'Shea DF. Synthetic applications of carbolithiation transformations. 2008.
- [5] Martinez-Estibalez U, Sotomayor N, Lete E. Intramolecular carbolithiation reactions for the synthesis of 2,4-disubstituted tetrahydro-quinolines: evaluation of TMEDA and (-)-sparteine as ligands in the stereoselectivity. *Org Lett*, 2009; 11: 1237-40.
- [6] Gomez-Sanjuan A, Sotomayor N, Lete E. Inter- and intramolecular enantioselective carbolithiation reactions. *Beilstein J Org Chem*, 2013; 9: 313-22.
- [7] Bahde RJ, Rychnovsky SD. Cyclization via Carbolithiation of  $\gamma$ -Amino Alkylolithium Reagents. *Org Lett*, 2008; 10: 4017-20.

- [8] Krief A, Remacle B, Mercier J. Synthesis of 1-Aryl-2-vinylcycloalkanes by Intramolecular Carbolithiation. *Synlett*, 2000; 2000: 1443-1446.
- [9] Liu T, Zhao X, Lu L, Cohen T. Readily Prepared 3-Chloro-1-(phenylthio)propene, a Versatile Three-Carbon Annulating Agent. 2009.
- [10] Fañanás FJ, Instituto Universitario de Química Organometálica “Enrique Moles” Unidad Asociada al C.S.I.C. JC, 8 Universidad de Oviedo, 33071 Oviedo (Spain) Fax: (+34) 98-5103450, Granados A, Instituto Universitario de Química Organometálica “Enrique Moles” Unidad Asociada al C.S.I.C. JC, 8 Universidad de Oviedo, 33071 Oviedo (Spain) Fax: (+34) 98-5103450, Sanz R, Departamento de Química Área de Química Orgánica Facultad de Ciencias UdBPMBsn, 09001-Burgos (Spain), Ignacio JM, Departamento de Química Área de Química Orgánica Facultad de Ciencias UdBPMBsn, 09001-Burgos (Spain), Barluenga J, Instituto Universitario de Química Organometálica “Enrique Moles” Unidad Asociada al C.S.I.C. JC, 8 Universidad de Oviedo, 33071 Oviedo (Spain) Fax: (+34) 98-5103450. Synthesis of Functionalized Pyrrole and Indole Derivatives through Carbometallation of Lithiated Double Bonds. *Chemistry - A European Journal*; 7: 2896-2907.
- [11] Sanz R, Ignacio JM, Rodriguez MA, Fananas FJ, Barluenga J. Intramolecular carbolithiation of 2,6-dilithio-1,6-heptadienes: an experimental and theoretical study. *Chemistry*, 2007; 13: 4998-5008.
- [12] Barluenga J, Fananas FJ, Sanz R, Marcos C. Intramolecular carbolithiation of allyl o-lithioaryl ethers: a new enantioselective synthesis of functionalized 2,3-dihydrobenzofurans. *Chemistry*, 2005; 11: 5397-407.
- [13] Gonzalez-Diaz H, Arrasate S, Gomez-San Juan A, Sotomayor N, Lete E, Besada-Porto L, Ruso JM. New Theory for Multiple Input-Output Perturbations in Complex Molecular Systems. 1. Linear QSPR Electronegativity Models in Physical, Organic, and Medicinal Chemistry. *Curr Top Med Chem*, 2013.
- [14] Gonzalez-Diaz H, Prado-Prado F, Ubeira FM. Predicting antimicrobial drugs and targets with the MARCH-INSIDE approach. *Curr Top Med Chem*, 2008; 8: 1676-90.
- [15] Gonzalez-Diaz H, Duardo-Sanchez A, Ubeira FM, Prado-Prado F, Perez-Montoto LG, Concu R, Podda G, Shen B. Review of MARCH-INSIDE & complex networks prediction of drugs: ADMET, anti-parasite activity, metabolizing enzymes and cardiotoxicity proteome biomarkers. In: ed.^eds., *Curr Drug Metab*: Netherlands, 2010; pp. 379-406.