J|A|C|S
ARTICLES
Published on Web 05/06/2006

Quantitative Structure−Activity Relationships of Ruthenium Catalysts for Olefin Metathesis

Giovanni Occhipinti, Hans-René Bjørsvik, and Vidar R. Jensen*

Contribution from the Department of Chemistry, University of Bergen, Allégaten 41, N-5007 Bergen, Norway

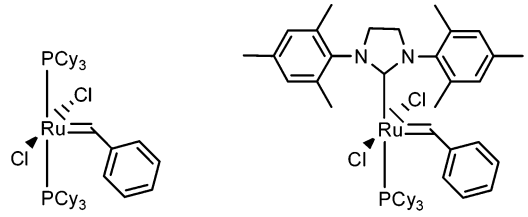
Received February 3, 2006; E-mail: Vidar.Jensen@kj.uib.no

Abstract: A quantitative structure−activity relationship (QSAR) model is presented in which both the independent and dependent (response) variables are derived from density functional theory (DFT) calculations on a large set of 14-electron complexes, $LCl_2Ru{=}CH_2$, with different dative ligands, L. The multivariate model thus correlates the properties of the 14-electron complexes with a calculated measure of activity, with modest computational cost, and reproduces the experimental order of activity for the Grubbs ruthenium catalysts for olefin metathesis. The accuracy and applicability of the model is to a large extent due to the use of highly specific geometric and electronic molecular descriptors which establish a direct connection between activity and chemically meaningful donor ligand properties. The ligands that most efficiently promote catalytic activity are those that stabilize the high-oxidation state (+4) metallacyclobutane intermediate relative to the ruthenium−carbene structures dominating the rest of the reaction pathway. Stabilization of the intermediate is ensured, among others, through ligand-to-metal σ donation, whereas metal-to-ligand π back-donation destabilizes the intermediate and lowers catalytic activity. A bulky dative ligand drives the reaction toward the less sterically congested metallacyclobutane species and thus contributes to catalytic activity. The multivariate model and the high-level descriptors furthermore provide practical handles for catalyst development as exemplified by the suggestion of several new donor ligands predicted to give more active and functional group tolerant ruthenium-based catalysts. The present strategy holds great promise for broader screenings of olefin metathesis catalysts as well as for development of homogeneous transition metal catalysts in general.

Introduction

Olefin metathesis has in recent years evolved into a very powerful and versatile tool in organic synthesis.[1,2] For example, starting from acyclic dienes, ring-closing metathesis (RCM) may be used in the synthesis of a large variety of carbo- and heterocycles and has proven extremely useful in the synthesis of natural products; see, e.g., ref 3. Olefin metathesis may also be applied in ring-opening polymerization (ROMP) to make functionalized polymers[4,5] that have not been achieved with conventional Ziegler-type catalysts despite decades of effort in the field of coordination polymerization. This remarkable progress has been spurred by the development of commercially available transition metal alkylidene catalysts for the olefin metathesis reaction,[6−8] notably the molybdenum or tungsten complexes of Schrock[6] and the ruthenium complexes developed by Grubbs.[7,8] While the Schrock catalysts display higher activities, the Grubbs family of catalysts (see Chart 1) have proven surprisingly air-stable and tolerant toward functional groups, and they react more readily with olefins than with acids, alcohols, and water.[8] The selectivity for the alkene double bond makes it possible to use readily available and inexpensive olefins

Chart 1. Grubbs-Type First (MC1)[7] and Second (MC2)[9,10] Generation Ruthenium-Based Catalysts for Olefin Metathesis

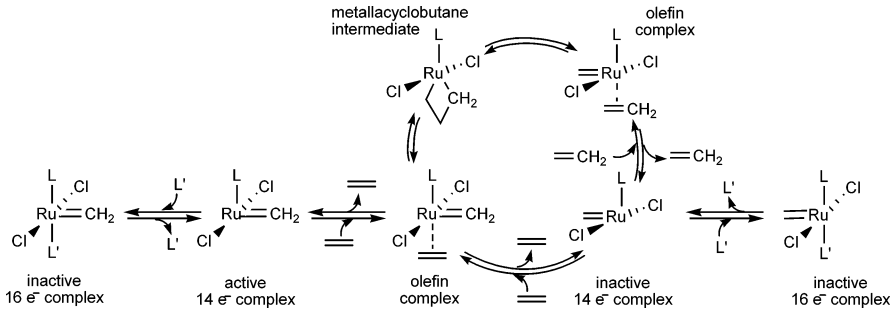


MC1 **MC2**

(1) Astruc, D. *New J. Chem.* **2005**, *29*, 42.
(2) Mori, M. *J. Synth. Org. Chem. Jpn.* **2005**, *63*, 423. Grubbs, R. H. *Tetrahedron* **2004**, *60*, 7117. *Handbook of Metathesis*; Grubbs, R. H., Ed.; Wiley-VCH: Weinheim, 2003. Fürstner, A. *Angew. Chem., Int. Ed.* **2000**, *39*, 3013.
(3) Fürstner, A.; Stelzer, F.; Rumbo, A.; Krause, H. *Chem.—Eur. J.* **2002**, *8*, 1856. Aissa, C.; Riveiros, R.; Ragot, J.; Fürstner, A. *J. Am. Chem. Soc.* **2003**, *125*, 15512. Fürstner, A. *Angew. Chem., Int. Ed.* **2003**, *42*, 3582. Fürstner, A.; Jeanjean, F.; Razon, P.; Wirtz, C.; Mynott, R. *Chem.—Eur. J.* **2003**, *9*, 307. Fürstner, A. *Eur. J. Org. Chem.* **2004**, 943. Nicolaou, K. C.; Vourloumis, D.; Winssinger, N.; Baran, P. S. *Angew. Chem., Int. Ed.* **2000**, *39*, 44.
(4) Scherman, O. A.; Kim, H. M.; Grubbs, R. H. *Macromolecules* **2002**, *35*, 5366. Bielawski, C. W.; Benitez, D.; Morita, T.; Grubbs, R. H. *Macromolecules* **2001**, *34*, 8610. Slugovc, C. *Macromol. Rapid Commun.* **2004**, *25*, 1283. Frenzel, U.; Nuyken, O. *J. Polym. Sci., Part A: Polym. Chem.* **2002**, *40*, 2895.
(5) Bielawski, C. W.; Grubbs, R. H. *Angew. Chem., Int. Ed.* **2000**, *39*, 2903.
(6) Schrock, R. R.; Murdzek, J. S.; Bazan, G. C.; Robbins, J.; Dimare, M.; Oregan, M. *J. Am. Chem. Soc.* **1990**, *112*, 3875. Schrock, R. R.; Hoveyda, A. H. *Angew. Chem., Int. Ed.* **2003**, *42*, 4592.
(7) Nguyen, S. T.; Johnson, L. K.; Grubbs, R. H.; Ziller, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 3974. Schwab, P.; France, M. B.; Ziller, J. W.; Grubbs, R. H. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2039.
(8) Trnka, T. M.; Grubbs, R. H. *Acc. Chem. Res.* **2001**, *34*, 18.

10.1021/ja060832i CCC: $33.50

Scheme 1. Hérisson−Chauvin Mechanism for Olefin Metathesis Adapted for the Grubbs Ruthenium Catalysts



as building blocks for unsaturated molecules and polymers that may otherwise be very difficult or even impossible to obtain. For these reasons, the ruthenium-based catalysts are the most widely used in organic synthesis, and much of the ongoing effort in catalyst development focuses on improving the Grubbs ruthenium catalysts.[1]

Detailed mechanistic and computational studies have been conducted parallel to the efforts in synthesis, and the resulting insight has, no doubt, inspired catalyst design and development in olefin metathesis. Molecular-level computational studies have contributed to the understanding of the formation and nature of the ruthenium alkylidene.[11,12] A series of experimental[12−16] and computational[16−24] studies have focused on the mechanism of the subsequent olefin metathesis process mediated by the ruthenium alkylidene complex or the decomposition of the latter.[25] It is now widely accepted that olefin metathesis mediated by the Grubbs-type ruthenium catalysts follows the general Hérisson−Chauvin mechanism which involves a metallacyclobutane intermediate.[26] The process is initiated by coordination of the olefin substrate onto a transition metal alkylidene (carbene). For the Grubbs-type ruthenium complexes, experimental[14,15,27] and computational studies[17,19,21,22] have shown that this olefin coordination is preceded by dissociation of a phosphine ligand, resulting in an overall mechanism as shown in Scheme 1. The lower activities for the Grubbs catalysts have been attributed to their mode of initiation.[28] Due to the expected trans-effect, the second dative ligand located opposite to the dissociating phosphine is arguably the single design element that has received the most attention in the efforts to develop more active Grubbs catalysts. A milestone in this optimization was reached with the introduction of complexes containing one N-heterocyclic carbene ligand (NHC) opposite to the dissociating phosphine, thus yielding the catalysts commonly referred to as the second-generation Grubbs ruthenium catalysts.[9,10]

Contemporary efforts to improve the initiation phase of the Grubbs family of catalysts involve the incorporation of the dissociating dative ligand into a loosely chelating group connected to the starting alkylidene ligand[29] or, alternatively, associating the phosphine ligand to the carbene by protonation of the latter.[28] Investigation of the effects from the remaining dative ligand has continued parallel to these developments, and synthesis of ruthenium alkylidene complexes containing new NHC and other donor ligands, followed by testing for catalytic activity and selectivity, constitutes an important line of research in the field of olefin metathesis (for recent reviews and examples, see refs 30 and 31). Clearly, knowledge of the relationship between the nature of the ligand L and the resulting

(9) Scholl, M.; Trnka, T. M.; Morgan, J. P.; Grubbs, R. H. *Tetrahedron Lett.* **1999**, *40*, 2247. Huang, J. K.; Stevens, E. D.; Nolan, S. P.; Petersen, J. L. *J. Am. Chem. Soc.* **1999**, *121*, 2674. Frenzel, U.; Weskamp, T.; Kohl, F. J.; Schattenman, W. C.; Nuyken, O.; Herrmann, W. A. *J. Organomet. Chem.* **1999**, *586*, 263. Ackermann, L.; Fürstner, A.; Weskamp, T.; Kohl, F. J.; Herrmann, W. A. *Tetrahedron Lett.* **1999**, *40*, 4787.
(10) Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H. *Org. Lett.* **1999**, *1*, 953.
(11) Spivak, G. J.; Coalter, J. N.; Olivan, M.; Eisenstein, O.; Caulton, K. G. *Organometallics* **1998**, *17*, 999. Coalter, J. N.; Spivak, G. J.; Gerard, H.; Clot, E.; Davidson, E. R.; Eisenstein, O.; Caulton, K. G. *J. Am. Chem. Soc.* **1998**, *120*, 9388. Coalter, J. N.; Bollinger, J. C.; Huffman, J. C.; Werner-Zwanziger, U.; Caulton, K. G.; Davidson, E. R.; Gerard, H.; Clot, E.; Eisenstein, O. *New J. Chem.* **2000**, *24*, 9. Ferrando-Miguel, G.; Coalter, J. N.; Gerard, H.; Huffman, J. C.; Eisenstein, O.; Caulton, K. G. *New J. Chem.* **2002**, *26*, 687. Dolker, N.; Frenking, G. *J. Organomet. Chem.* **2001**, *617*, 225. Bernardi, F.; Bottoni, A.; Miscione, G. P. *Organometallics* **2000**, *19*, 5529.
(12) Hansen, S. M.; Rominger, F.; Metz, M.; Hofmann, P. *Chem.−Eur. J.* **1999**, *5*, 557. Volland, M. A. O.; Hansen, S. M.; Hofmann, P. In *Chemistry at the Beginning of the Third Millennium: Molecular Design, Supramolecules, Nanotechnology and Beyond*; Fabrizzi, L., Poggi, A., Eds.; Springer: Berlin, 2000.
(13) Sanford, M. S.; Ulman, M.; Grubbs, R. H. *J. Am. Chem. Soc.* **2001**, *123*, 749. Sanford, M. S.; Love, J. A.; Grubbs, R. H. *J. Am. Chem. Soc.* **2001**, *123*, 6543. Ulman, M.; Grubbs, R. H. *Organometallics* **1998**, *17*, 2484. Love, J. A.; Sanford, M. S.; Day, M. W.; Grubbs, R. H. *J. Am. Chem. Soc.* **2003**, *125*, 10103. Adlhart, C.; Chen, P. *Helv. Chim. Acta* **2000**, *83*, 2192. Adlhart, C.; Volland, M. A. O.; Hofmann, P.; Chen, P. *Helv. Chim. Acta* **2000**, *83*, 3306. Adlhart, C.; Chen, P. *Helv. Chim. Acta* **2003**, *86*, 941. Basu, K.; Cabral, J. A.; Paquette, L. A. *Tetrahedron Lett.* **2002**, *43*, 5453. Lehman, S. E.; Wagener, K. B. *Macromolecules* **2002**, *35*, 48. Stuer, W.; Wolf, J.; Werner, H. *J. Organomet. Chem.* **2002**, *641*, 203. Volland, M. A. O.; Adlhart, C.; Kiener, C. A.; Chen, P.; Hofmann, P. *Chem.−Eur. J.* **2001**, *7*, 4621. Hofmann, M.; Puskas, J. E.; Weiss, K. *Eur. Polym. J.* **2002**, *38*, 19.
(14) Dias, E. L.; Nguyen, S. T.; Grubbs, R. H. *J. Am. Chem. Soc.* **1997**, *119*, 3887.
(15) Hinderling, C.; Adlhart, C.; Chen, P. *Angew. Chem., Int. Ed.* **1998**, *37*, 2685.
(16) Adlhart, C.; Hinderling, C.; Baumann, H.; Chen, P. *J. Am. Chem. Soc.* **2000**, *122*, 8204.
(17) Aagaard, O. M.; Meier, R. J.; Buda, F. *J. Am. Chem. Soc.* **1998**, *120*, 7174. Fomine, S.; Vargas, S. M.; Tlenkopatchev, M. A. *Organometallics* **2003**, *22*, 93. Bernardi, F.; Bottoni, A.; Miscione, G. P. *Organometallics* **2003**, *22*, 940.
(18) Meier, R. J.; Aagaard, O. M.; Buda, F. *J. Mol. Catal. A: Chem.* **2000**, *160*, 189. Costabile, C.; Cavallo, L. *J. Am. Chem. Soc.* **2004**, *126*, 9592. Vyboishchikov, S. E.; Thiel, W. *Chem.−Eur. J.* **2005**, *11*, 3921. Fomine, S.; Ortega, J. V.; Tlenkopatchev, M. A. *J. Mol. Catal. A: Chem.* **2005**, *236*, 156. Lippstreu, J. J.; Straub, B. F. *J. Am. Chem. Soc.* **2005**, *127*, 7444.
(19) Vyboishchikov, S. E.; Bühl, M.; Thiel, W. *Chem.−Eur. J.* **2002**, *8*, 3962.
(20) Adlhart, C.; Chen, P. *Angew. Chem., Int. Ed.* **2002**, *41*, 4484.
(21) Cavallo, L. *J. Am. Chem. Soc.* **2002**, *124*, 8965.
(22) Adlhart, C.; Chen, P. *J. Am. Chem. Soc.* **2004**, *126*, 3496.
(23) Straub, B. F. *Angew. Chem., Int. Ed.* **2005**, *44*, 5974.
(24) Tsipis, A. C.; Orpen, A. G.; Harvey, J. N. *Dalton Trans.* **2005**, 2849.
(25) van Rensburg, W. J.; Steynberg, P. J.; Meyer, W. H.; Kirk, M. M.; Forman, G. S. *J. Am. Chem. Soc.* **2004**, *126*, 14332.
(26) Hérisson, J.-L.; Chauvin, Y. *Makromol. Chem.* **1970**, *141*, 161.
(27) Tallarico, J. A.; Bonitatebus, P. J.; Snapper, M. L. *J. Am. Chem. Soc.* **1997**, *119*, 7157.
(28) Romero, P. E.; Piers, W. E.; McDonald, R. *Angew. Chem., Int. Ed.* **2004**, *43*, 6161.
(29) Garber, S. B.; Kingsbury, J. S.; Gray, B. L.; Hoveyda, A. H. *J. Am. Chem. Soc.* **2000**, *122*, 8168. Kingsbury, J. S.; Harrity, J. P. A.; Bonitatebus, P. J.; Hoveyda, A. H. *J. Am. Chem. Soc.* **1999**, *121*, 791. Zaja, M.; Connon, S. J.; Dunne, A. M.; Rivard, M.; Buschmann, N.; Jiricek, J.; Blechert, S. *Tetrahedron* **2003**, *59*, 6545. Wakamatsu, H.; Blechert, S. *Angew. Chem., Int. Ed.* **2002**, *41*, 2403. Wakamatsu, H.; Blechert, S. *Angew. Chem., Int. Ed.* **2002**, *41*, 794. Grela, K.; Harutyunyan, S.; Michrowska, A. *Angew. Chem., Int. Ed.* **2002**, *41*, 4038. Michrowska, A.; Bujok, R.; Harutyunyan, S.; Sashuk, V.; Dolgonos, G.; Grela, K. *J. Am. Chem. Soc.* **2004**, *126*, 9318.

activity (or selectivity) is vital in these efforts in order to subject to synthesis only those ligands and complexes that stand good chances of giving active and selective catalysts. To date, researchers have used their general chemical knowledge and intuition or drawn upon explicit experimental or computational data available for the ruthenium catalysts in order to make "ad hoc", qualified guesses for their lead structures. However, there exist strategies in order to systemize such relationships and automate their use in the generation of lead structures. Quantitative structure−property and structure−activity relationships (QSPR/QSAR) based on theoretically generated molecular descriptors[32] have so far attracted only little attention in organometallic chemistry and catalysis. Such techniques, however, are recognized as standard and important in drug design.[33] Recent reports indicate that statistical treatment of theoretical molecular descriptors may be useful for characterization and prediction of ligands in organometallic chemistry[34] and also that a potential for computer-aided "catalyst design" should exist in chemistry.[35] In the current contribution we attempt to realize some of this potential by proposing and testing procedures in order to establish a more precise and systematic relationship between the nature of the design element, L, and the catalytic activity of the catalyst $LL'Cl_2Ru{=}CH_2$. The goal is to contribute to a more cost-efficient optimization of the Grubbs family of catalysts.

To this end we have constructed QSPR/QSAR models that correlate the structures and properties of a large set of 14-electron ruthenium alkylidene complexes, $LCl_2Ru{=}CH_2$, with the activities for a smaller, representative set of catalysts. We use a protocol based on multivariate statistical and mathematical methods for which both the descriptors (forming the independent variables) and the activities (the response variables) are derived from density functional theory (DFT) calculations. The multivariate models offer unique insight into the steric and electronic effects governing the activity of the Grubbs family of ruthenium-based catalysts for olefin metathesis. This insight enables us to answer questions of a recent debate, e.g., whether steric pressure on the alkylidene moiety may favor formation of the metallacyclobutane intermediate, thereby lowering the barrier to olefin metathesis.[21,22] Finally, the multivariate models have been used to predict the activities of new ruthenium alkylidene complexes, some of which are suggested to be more active than the most active existing second-generation Grubbs catalysts.

Computational Details

Density Functional Theory (DFT) Calculations. Initial conformational searches were performed using a semiempirical method (PM3) implemented in Spartan'02.[36] The thus located low-energy conformers were taken as input structures for geometry optimization in the Gaussian 03[37] suite of programs using the OLYP density functional. All located stationary points were characterized by calculation of the Hessian matrix. The OLYP functional contains Handy's OPTX modification[38] of Becke's exchange[39] and correlation due to Lee, Yang, and Parr.[40] This functional has been reported to be superior to the related BLYP functional and other pure density functionals in general[41] and also in applications involving transition metals.[42] Numerical integrations were performed using the default "fine" grid of Gaussian 03, and the Gaussian 03 default values were chosen for the self-consistent-field (SCF) and geometry optimization convergence criteria. Thermochemical values were computed within the harmonic-oscillator, rigid-rotor, and ideal-gas approximations.

The basis sets used in the geometry optimizations were the built-in "LANL2DZ" basis sets of Gaussian 03, which imply the use of Hay and Wadt effective core potentials (ECPs) in combination with their accompanying valence double-ζ basis sets for all elements beyond the first row.[43,44] For ruthenium, the ECP replaced the 1s, 2s, 2p, 3s, 3p, and 3d electrons, whereas the 4s, 4p, 4d, 5s, and 5p orbitals were represented by the Hay and Wadt primitive basis set[44] (5s,6p,4d) contracted to [3s,3p,2d]. First row and hydrogen atoms were described by standard Dunning and Hay valence double-ζ basis sets.[45]

Total energies and properties were obtained in single-point (SP) energy evaluations using the three-parameter hybrid density functional method of Becke (termed "B3LYP"),[46] as implemented in the Gaussian 03 set of programs.[37] The SP calculations involved basis sets that were improved compared to those used in the geometry optimizations: For ruthenium, the Hay and Wadt primitive basis set,[44] (5s,6p,4d), was contracted to [4s,4p,3d]. A diffuse s function was added to the basis sets of all other elements, and a diffuse p function was added to all non-hydrogen elements (except Ru). The diffuse s functions of the non-hydrogen elements were added in an even-tempered manner, whereas the s exponent for hydrogen and the p exponents were taken from ref 47. For hydrogen, a p polarization function was added, whereas a d polarization function was added to the basis sets of all other elements (except Ru).[47] These diffuse and polarization functions were added to the standard valence double-ζ basis sets described above and contracted to [4s,1p] for hydrogen and [4s,4p,1d] for first-row elements as well as for the valence of elements beyond the first row (except Ru).

(30) Despagnet-Ayoub, E.; Grubbs, R. H. *Organometallics* **2005**, *24*, 338. Forman, G. S.; McConnell, A. E.; Hanton, M. J.; Slawin, A. M. Z.; Tooze, R. P.; van Rensburg, W. J.; Meyer, W. H.; Dwyer, C.; Kirk, M. M.; Serfontein, D. W. *Organometallics* **2004**, *23*, 4824. Pruhs, S.; Lehmann, C. W.; Fürstner, A. *Organometallics* **2004**, *23*, 280. Opstal, T.; Verpoort, F. *J. Mol. Catal. A: Chem.* **2003**, *200*, 49. Dinger, M. B.; Mol, J. C. *Adv. Synth. Catal.* **2002**, *344*, 671.

(31) Trnka, T. M.; Morgan, J. P.; Sanford, M. S.; Wilhelm, T. E.; Scholl, M.; Choi, T. L.; Ding, S.; Day, M. W.; Grubbs, R. H. *J. Am. Chem. Soc.* **2003**, *125*, 2546.

(32) Loew, G. H.; Villar, H. O.; Alkorta, I. *Pharm. Res.* **1993**, *10*, 475. Karelson, M.; Lobanov, V. S.; Katritzky, A. R. *Chem. Rev.* **1996**, *96*, 1027. Apostolakis, J.; Caflisch, A. *Comb. Chem. High Throughput Screening* **1999**, *2*, 91. Carbo-Dorca, R.; Amat, L.; Besalu, E.; Girones, X.; Robert, D. *THEOCHEM* **2000**, *504*, 181. de Julian-Ortiz, J. V. *Comb. Chem. High Throughput Screening* **2001**, *4*, 295. Bajorath, F. *Nat. Rev. Drug Discovery* **2002**, *1*, 882.

(33) Weng, Z. P.; DeLisi, C. *Trends Biotechnol.* **2002**, *20*, 29. Manly, C. J.; Louise-May, S.; Hammer, J. D. *Drug Discovery Today* **2001**, *6*, 1101. Fenniri, H. *Curr. Med. Chem.* **1996**, *3*, 343.

(34) Fey, N.; Tsipis, A. C.; Harris, S. E.; Harvey, J. N.; Orpen, A. G.; Mansson, R. A. *Chem.−Eur. J.* **2005**, *12*, 291. Cooney, K. D.; Cundari, T. R.; Hoffman, N. W.; Pittard, K. A.; Temple, M. D.; Zhao, Y. *J. Am. Chem. Soc.* **2003**, *125*, 4318. Perrin, L.; Clot, E.; Eisenstein, O.; Loch, J.; Crabtree, R. H. *Inorg. Chem.* **2001**, *40*, 5806. Bosque, R.; Sales, J. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 225.

(35) Cruz, V. L.; Ramos, J.; Martinez, S.; Munoz-Escalona, A.; Martinez-Salazar, J. *Organometallics* **2005**, *24*, 5095. Burello, E.; Farrusseng, D.; Rothenberg, G. *Adv. Synth. Catal.* **2004**, *346*, 1844. Lipkowitz, K. B.; Kozlowski, M. C. *Synlett* **2003**, 1547. Lu, Q. Z.; Yu, R. Q.; Shen, G. L. *J. Mol. Catal. A: Chem.* **2003**, *198*, 9. Arnold, F. H. *Nature* **2001**, *409*, 253.

(36) *Spartan'02*; Wavefunction, Inc.: Irvine, CA, 2002.

(37) Frisch, M. J. et al. *Gaussian 03*, revision B.04; Gaussian, Inc.: Pittsburgh, PA, 2003.

(38) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403.

(39) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(40) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(41) Xu, X.; Goddard, W. A. *J. Phys. Chem. A* **2004**, *108*, 8495. Baker, J.; Pulay, P. *J. Chem. Phys.* **2002**, *117*, 1441. Hoe, W. M.; Cohen, A. J.; Handy, N. C. *Chem. Phys. Lett.* **2001**, *341*, 319.

(42) Baker, J.; Pulay, P. *J. Comput. Chem.* **2003**, *24*, 1184.

(43) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 284.

(44) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.

(45) Dunning, T. H., Jr.; Hay, P. J. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum Press: New York, 1977; p 1.

(46) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(47) Check, C. E.; Faust, T. O.; Bailey, J. M.; Wright, B. J.; Gilbert, T. M.; Sunderlin, L. S. *J. Phys. Chem. A* **2001**, *105*, 8111.

Multivariate Modeling. The multivariate regression and data analyses were carried out by means of the PLS Toolbox 3.5[48] written for MATLAB.[49] The variables were molecular descriptors calculated for an initial set of 82 active complexes, $LCl_2Ru{=}CH_2$ (see Chart 2). A large number of these descriptors (308) were standard descriptors (constitutional, topological, etc.) calculated by the Codessa 2.62 program[50] on the basis of the geometry and (for some descriptors) properties from the DFT calculations. A set of additional geometric or electronic descriptors (35) were extracted or calculated from the DFT properties using software developed in-house, among others with the goal to represent the charge distribution (electrostatic potential) of the catalyst complex. Our active complexes have a common identical fragment, $Cl_2Ru{=}CH_2$, for which atomic properties such as charges can be included as descriptors. To describe the molecular charge distribution we have thus included atomic charges calculated according to the CHELPG scheme.[51] These charges are fitted to reproduce as closely as possible the electrostatic field generated by the charge distribution of the molecule. The exclusion radii, within which the electrostatic field is not evaluated, were set to 2.00 Å for Ru, 2.40 Å for Sn, 2.20 Å for As, and 2.30 Å for Se, whereas Gaussian 03 default values were used for all other elements. CHELPG atomic charges for ruthenium and the alkylidene carbon atom were included as molecular descriptors in addition to the average atomic charge for the two equatorial chlorine atoms and the group charge of the CH_2 alkylidene. In addition, the calculated charge on the fragment, $Cl_2Ru{=}CH_2$, was included and used as an estimate of the net L→Ru donation. Moreover, back-donation from the ruthenium d_π orbitals to the dative ligand was approximated as the difference between the occupation of the two Ru lone pairs (d_{xz} and d_{yz} with the molecule oriented with the Ru−L bond along the *z*-axis) in the $Cl_2Ru{=}CH_2$ fragment alone (3.96) and the corresponding occupation in each of the $LCl_2Ru{=}CH_2$ complexes, as obtained using natural bond orbital (NBO) analysis.[52,53] This furthermore allowed us to include estimates of the L→Ru σ donation as the difference between the net L→Ru donation and the Ru d_π→L_π back-donation. Steric exchange interactions between the ligand, L, and the CH_2 moiety[54] were calculated using natural steric analysis.[55]

All descriptors were mean centered and scaled to unit variance because the numerical values of the descriptors vary significantly. To minimize subsequent problems of chance correlation, we started out by pruning the 342 descriptors calculated initially. Descriptors having the same value for more than 50% of the complexes (117), and most of the strongly correlated (redundant) descriptors (104) were removed. A complete list of the remaining descriptors (122) can be found in the Supporting Information. Hotelling's T^2 statistics[56] revealed the presence of five outliers, the complexes of ligands **O1**, **C2**, **A30**, **A36**, and **A7** in the initial set of 82 metal complexes (see Chart 2 for a complete list of ligands), and complexes of these ligands were thus removed from the set, leaving 77 complexes for further analysis.

A representative subset of 33 alkylidene complexes, $Cl_2LRu{=}CH_2$, for which response variables were calculated, were selected by k-means nearest group cluster analysis using the Mahalanobis distance[57] and principal component analysis (PCA)[58] in a model retaining eight principal components which explained 79% of the variance. The cluster analysis is shown in the Supporting Information. The representative complexes formed the calibration set of our QSAR models. To test the predictive ability of these models, nine of the complexes in the initial set were randomly selected to form a test set.

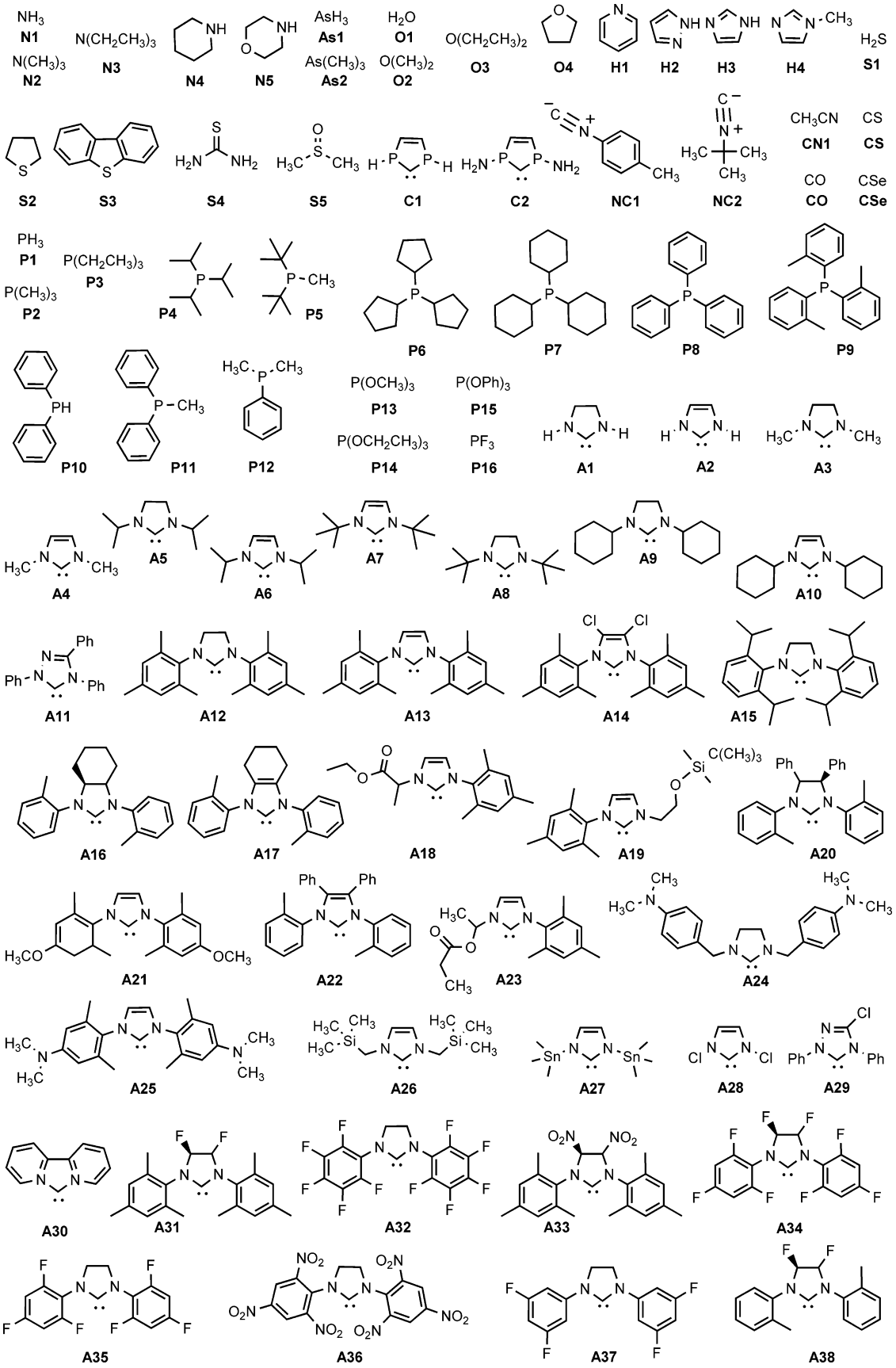
Correlating the independent variables, the descriptors, with experimentally recorded catalyst activities is difficult. Existing studies of activity do not encompass broad ranges of catalysts, and variation in olefin substrate and experimental conditions preclude simultaneous use of several different experimental studies in the multivariate modeling. Our response variables were thus calculated DFT energy differences, termed "productivities". These define the balance between 16-electron precursor-type complexes, $LL'Cl_2Ru{=}CH_2$, and metallacyclobutane intermediates as defined in Figure 1. The energy of the precursor-type complexes is taken as the mean value for L′ = PMe_3, H_2O, and CH_2O. PMe_3 is a model of the phosphine ligand dissociating from the precursor (see Scheme 1), usually PCy_3, whereas water and formaldehyde are models of oxygen-containing functional groups, present for example in the olefin substrates, that may bind to and inactivate the catalyst. Our energy reference point thus should offer a practical measure of the likeliness that the active complexes, $LCl_2Ru{=}CH_2$, are inactivated by complexation of Lewis bases, L′. The calculated productivity of a given complex thus defines the balance between several possible inactivated states and a key intermediate of the olefin metathesis reaction, the metallacyclobutane complex, and is related to both activity and functional group tolerance. The calculated relative energies in a series of quantum chemical studies suggest that the stability of the metallacyclobutane intermediate correlates with the barrier height of the olefin metathesis reaction, regardless of the exact nature of the barrier.[19−22] For example, Table 1 of ref 22 offers a comparison of energies of the metallacyclobutane intermediates and the corresponding energies of the transition states for their formation covering 10 different ruthenium Grubbs-type complexes. Simple linear regression between the stabilities of the metallacyclobutane intermediates, on one hand, and the reaction barriers (relative to the 16-electron precursor complexes), on the other hand, gives a correlation coefficient, $R > 0.99$. The stability of the key intermediate, the metallacyclobutane, thus seems to offer a useful approximate measure for the relative barrier heights when comparing different olefin metathesis catalysts, thereby avoiding the more costly optimization of the transition states. Finally, the productivities calculated for the representative subset of complexes, forming a $\mathbf{y}_{33} \times 1$ response vector, were correlated to the molecular descriptors, forming an $\mathbf{X}_{33} \times 122$ predictor matrix, using partial least squares regression (PLSR).[58]

To assess the quality of this multivariate model, leave-one-out cross validation[58] and prediction of the test set were performed. The root-mean-square error of cross validation, RMSECV,[58] and the root-mean-square error of prediction, RMSEP,[58] of the nine complexes of the test set were compared with the calibration range for the productivity (29.4 kcal/mol, from −23.8 to +5.6 kcal/mol).

In this model, the first eight PLSR components (factors) explain 99.92% of the variance in **y**, with the ninth component adding only 0.03%. With eight factors, the corresponding RMSECV is 1.72 kcal/mol and the RMSEP is 0.83 kcal/mol, corresponding to 5.9% and 2.8% of the calibration gap, respectively. We proceeded to improve the model

(48) Wise, B. M.; Gallagher, N. B.; Bro, R.; Shaver, J. M.; Windig, W.; Koch, R. S. *PLS Toolbox*, version 3.5; Eigenvector Technologies: Manson, WA, 2004.
(49) *MATLAB*, version 7.0.4; The MathWorks Inc.: Natick, MA, 2004.
(50) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *Comprehensive Descriptors for Structural and Statistical Analysis (Codessa)*, version 2.62; Semichem, Inc.: Shawnee Mission, KS, 2004.
(51) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
(52) Glendening, E. D.; J., K. B.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Weinhold, F. *NBO*, version 5.0.; Theoretical Chemistry Institute, University of Wisconsin, 2001.
(53) The approximation of donation and back-donation by differences in orbital populations and atomic partial charges is based on the assumption that effects from geometry relaxation and intrafragment polarization remain essentially constant between the different ligands. For all the complexes, the Lewis structure specified in the NBO calculations contained a $Ru{=}CH_2$ double bond in addition to two Ru−Cl single bonds and one Ru−L single bond.
(54) The CH_2 moiety employed in the calculations of steric exchange repulsion in Ru alkylidene complexes included the occupied valence orbitals of the carbene carbon atom, the C−H bonds, and the metal−carbon σ- and π-bonds. In the metallacyclobutane intermediate, the corresponding σ-bonded CH_2 group was used in addition to the second $Ru{-}CH_2$ σ-bond in the metallacycle, resulting in a fragment including the same number of electrons as the CH_2 moiety defined for the Ru alkylidene complexes. For the ligand L, all the valence electrons were used, including those of the Ru−L bond.
(55) Badenhoop, J. K.; Weinhold, F. *J. Chem. Phys.* **1997**, *107*, 5422.
(56) Winer, B. J.; Brown, D. R.; Michels, K. M. *Statistical principles in experimental design*, 3rd ed.; McGraw-Hill: New York, 1991.
(57) Mahalanobis, P. C. *Proc. Natl. Inst. Sci. India* **1936**, *2*, 49.
(58) Malinowski, E. R. *Factor analysis in chemistry*, 3rd ed.; Wiley: New York, 2002.

Chart 2. Ligands, L, Used in the Initial Set of 82 Active Complexes, $LCl_2Ru{=}CH_2$



by performing a manual pruning of descriptors based on their relative contribution in the model and retained 22 molecular descriptors in the final PLSR model. The first four factors now explain 99.58% of the variance in **y**, and the fifth factor adds only 0.07%. With four factors, the corresponding RMSECV is 0.77 kcal/mol, and the RMSEP, 0.53 kcal/mol, corresponding to 2.6% and 1.8% of the calibration range,

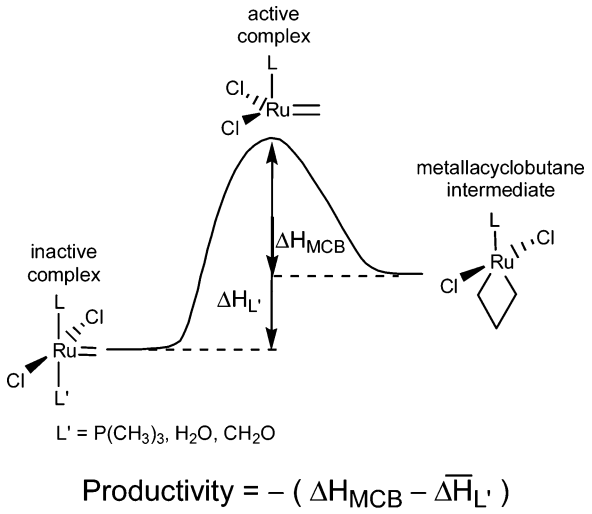


Figure 1. Definition of "productivity" as used as the response variable in the QSAR model building. The two enthalpy differences are given with respect to the 14-electron active complex. $\Delta H_L'$ is thus always negative, and ΔH_{MCB} is found to be positive only for a few of the catalysts with very low calculated productivities.

respectively. This new and smaller model thus has a better predictive ability than the initial model based on 122 descriptors and performs excellently given the broad range of complexes present in our set. The model is presented in detail in the following, with a focus on chemically significant structure−activity relationships and prediction of new catalysts.

Results and Discussion

Validation of Multivariate Model. The present study covers a broad range of ligands for organometallic chemistry, from the small and π acidic CO molecule to contemporary and sterically demanding N-heterocyclic carbenes for which virtually no π back-donation is expected. A complete list of the 82 ligands initially included (of which five were discarded as outliers) is given in Chart 2. A representative subset (33) of complexes were selected by cluster analysis (see the Computational Details section for more information), and the productivities (for definition, see Figure 1) of these were calculated explicitly. The predictions of the PLSR model ensuring maximum covariance between the calculated productivities and the corresponding molecular descriptors are shown in Figure 2. The predicted and explicitly calculated productivities for all the complexes can be found in the Supporting Information. The span in productivity, almost 30 kcal/mol, shows that the neutral ligand has a remarkable influence on the stability of the metallacyclobutane intermediate. The variation in productivity furthermore confirms our expectation that there are large differences among the current ligands. Nevertheless, the PLSR model performs well, with errors on the order of 1 kcal/mol. The root-mean-square error of cross validation, RMSECV, is 0.77 kcal/mol, and the root-mean-square error of prediction, RMSEP, 0.53 kcal/mol. The latter error implies that the productivities for the nine complexes of the test set (green bars in Figure 2) can be predicted with an error amounting to 1.8% of the calibration range. In other words, the multivariate model predicts the productivity accurately and thus offers a cost-effective means of obtaining this energy difference for a given candidate catalyst by a single calculation on the corresponding 14-electron complex.

Having established that the multivariate model is able to produce reasonably accurate productivities, we now turn to the question of whether the productivities reflect the experimentally recorded activities, i.e., whether productivity, as defined in Figure 1, appears to be a good approximation to catalytic activity. First, it can be noticed that phosphine complexes generally have lower productivities than complexes of N-heterocyclic carbenes (NHCs), thus reproducing the observation that second-generation Grubbs-type catalysts are more active than the first-generation catalysts. Furthermore, **P7** (PCy_3) is predicted to give a more active catalyst (predicted productivity, $P = -3.5$ kcal/mol) than **P4** (P^iPr_3, $P = -5.5$ kcal/mol) and significantly better than **P8** (PPh_3, $P = -10.7$ kcal/mol), reflecting the observed[8,14] order of (decreasing) activity for the phosphines: $PCy_3 > P^iPr_3 \gg PPh_3$. Phosphine **P6**, which is observed to give a catalyst with an activity comparable to that of **P4**,[59] is predicted to give a slightly more active catalyst ($P = -3.2$ kcal/mol) than **P7**, and this discrepancy probably reflects a minor inaccuracy in the multivariate model. Turning to the predicted trends in activity for the NHC ligands, we obtain the following productivities for a few of the carbenes most commonly used in the second-generation Grubbs catalysts: **A14** (Cl_2IMes) 1.1 kcal/mol, **A13** (IMes) 2.3 kcal/mol, **A12** (H_2IMes) 2.4 kcal/mol. The calculated productivities correspond to the experimentally observed[31] order of (increasing) activities for the catalysts: $(Cl_2IMes)(PCy_3)(Cl)_2RuCHPh$ < $(IMes)(PCy_3)(Cl)_2$-$RuCHPh$ < $(H_2IMes)(PCy_3)(Cl)_2RuCHPh$.

The predicted productivity of the catalyst based on the H_2IMes ligand is 0.1 kcal/mol higher than that of the IMes ligand, whereas the corresponding explicitly calculated productivities are less similar, 2.6 and 1.8 kcal/mol, for the catalyst based on the saturated and unsaturated ligand, respectively. Our results thus consistently predict that H_2IMes should give a more active catalyst than IMes, in agreement with experimental observation.[5,10,31,60,61]

A carbene ligand with *tert*-butyl substituents on the nitrogen atoms (**A8**) is predicted to give the most active catalyst of the ones studied here (Figure 2). Unfortunately, the resulting steric effects weaken the bond to the metal,[62] and the weak Ru−L bond is probably the reason for the low stability (and activity) noted for a closely related catalyst (**A7**) that has actually been synthesized.[63] In conclusion, a good qualitative agreement between our theoretically predicted activities and those observed experimentally seems to be established. A more detailed comparison of experimental and theoretical activities is difficult since the experimental activities are very dependent on the substrate alkene and the type of metathesis reaction.

Structure−Activity Relationships

Electronic Effects. The multivariate model offers insight into the electronic and steric factors governing activity and functional group tolerance. The molecular descriptors and their contributions in the final QSAR model are given in Table 1, and the complete correlation map between the descriptors and the

(59) Schwab, P.; Grubbs, R. H.; Ziller, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 100.
(60) Schramm, M. P.; Reddy, D. S.; Kozmin, S. A. *Angew. Chem., Int. Ed.* **2001**, *40*, 4274.
(61) Hillier, A. C.; Sommer, W. J.; Yong, B. S.; Petersen, J. L.; Cavallo, L.; Nolan, S. P. *Organometallics* **2003**, *22*, 4322.
(62) Ligand **A8** has significant steric repulsion toward the rest of the catalyst, $Cl_2Ru{=}CH_2$, and a calculated bond dissociation enthalpy ($\Delta H_{298} = 36.4$ kcal/mol) 15−20 kcal/mol lower than those of contemporary carbene ligands used in Grubbs ruthenium catalysts for olefin metathesis.
(63) Weskamp, T. PhD Thesis, Technische Universität München, Munich, 1999.

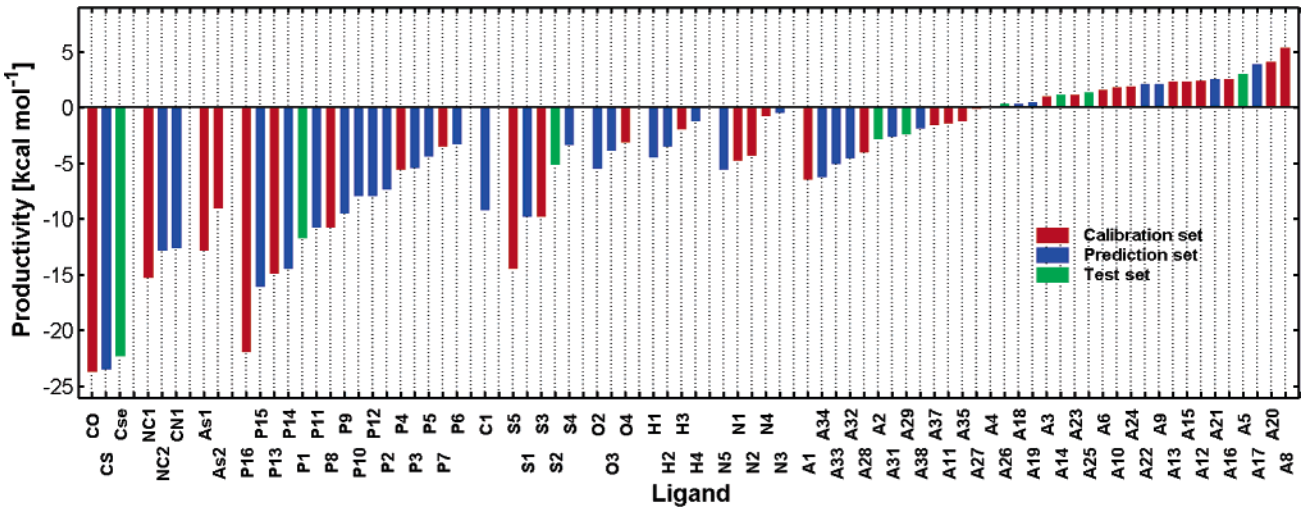


Figure 2. Predicted productivity of catalysts $LCl_2Ru{=}CH_2$ for all ligands, L. The structures of the ligands are shown in Chart 2. A table containing all the predicted and explicitly calculated productivities is given in the Supporting Information.

response, the productivity (defined in Figure 1), is given in Figure 3. The electronic descriptor which is most strongly correlated with productivity is the Wiberg bond order index[64] for the $Ru{=}CH_2$ bond (descriptor no. 10). A high bond order (BO) is positively correlated, whereas the corresponding bond distance (descriptor no. 11) is negatively correlated with productivity. Both descriptors thus show that a strong ruthenium−alkylidene bond is important for catalytic activity and the highest productivities are obtained for catalysts displaying a $Ru{=}CH_2$ Wiberg bond order well above 1.6. The best Lewis structures obtained in resonance structure analysis performed with the NBO program[52] contained a $Ru{=}CH_2$ double bond, and two natural bonding orbitals are always found between ruthenium and the alkylidene carbon which carries a partial negative charge (CHELPG[51]) in the range −0.1 *e* to −0.5 *e*. The bulk of the present carbenes can thus be characterized as having more Schrock than Fischer character. This furthermore indicates that they should rather be considered as complexes of ruthenium(+4) although they usually are referred to as ruthenium(+2) complexes, i.e., carbene complexes of ruthenium in oxidation state +2. In the metallacyclobutane intermediate, however, ruthenium has oxidation state +4. Thus, the strong correlation between $Ru{=}CH_2$ bond strength and productivity indicates that the 14-electron complexes which are already effectively in a high oxidation state should form comparably stable metallacyclobutane intermediates; i.e., these complexes need less activation in order to increase their oxidation state. It should be noted that the variation in BO and bond distance for the $Ru{=}CH_2$ bond in the 14-electron complexes does not originate from different conformers of the alkylidene group. The conformation of the alkylidene is identical in all the 14-electron complexes, with the CH_2-plane being oriented essentially orthogonal to the Cl_2Ru-plane.

L→Ru σ-donation (descriptor no. 22) also has a significant (positive) contribution in the multivariate model (see Table 1). The clear relationship between the donating abilities of the dative ligand and catalytic performance (productivity) is also evident in Table 2 where L→Ru σ-donation is seen to follow closely the experimentally recorded trend in catalytic activity, e.g., with σ-donation for the N-heterocyclic carbene ligands of the Grubbs second-generation catalysts estimated to approach 0.5 *e*, more than 0.1 *e* larger than for the most basic phosphines of the first-generation catalysts. The hierarchy of σ-donating capabilities obtained for the ligands in the present work is as expected, with PF_3 and H_2IMes defining the two extremes among the selection of ligands in Table 2. The latter ligand is calculated to be only a marginally better σ-donor than the corresponding unsaturated ligand, IMes, in accord with recent results based on calculations and measurements of Ru−L bond dissociation energies.[61] Almost identical Ru−carbene bond mechanisms have been noted for saturated and unsaturated Arduengo carbenes in a recent combined experimental and computational study.[65]

σ-Donation from the ligand correlates with productivity as well as a series of other electronic descriptors (see the correlation map, Figure 3). An obvious correlation is seen to exist between σ-donation and the dipole moment of the complex (descriptor no. 16). The resulting increase in electron density on the metal, for example, is followed by a reduction in the Fukui atomic electrophilic reactivity index for the Ru atom (descriptor no. 6), implying that complexes with less electrophilic ruthenium atoms give the more active catalysts. The Wiberg BO (positive correlation) and the bond distance (negative correlation) of the ruthenium−alkylidene bond both also show moderate covariance with the L→Ru σ-donation, suggesting that some of the increased ruthenium−alkylidene bond strength is ensured through the added electron density resulting from σ-donation from the dative ligand.

Part of the electron density donated from L is absorbed by the chlorines, making these atoms increasingly negatively charged and the Ru−Cl bonds more ionic. The average charge on the chlorine atoms correlates negatively with productivity (the correlation coefficient, $R = -0.84$), meaning that large negative partial charges (close to −0.5 *e*) on the chlorines are found in the catalysts of high productivity. The corresponding Ru−Cl bonds become less covalent (lower BO) and longer, and because they form wide angles (> 90°) with the $Ru{=}CH_2$ bond in the 14-electron active complexes, the result is a weaker trans influence on the latter bond. The overall result of more ionic Ru−Cl bonds and higher bond orders for the $Ru{=}CH_2$ bond is

(64) Wiberg, K. B. *Tetrahedron* **1968**, *24*, 1083.

(65) Ho, V. M.; Watson, L. A.; Huffman, J. C.; Caulton, K. G. *New J. Chem.* **2003**, *27*, 1446.

Table 1. Molecular Descriptors and Their Regression Coefficients (β) in the QSAR Model

descriptor number	molecular descriptor	β^m
11	Ru=CH_2 bond distance	−1.419
2	Kier and Hall index (order 2)[a]	−0.985
18	max bond order of a C atom	−0.730
6	electrophilic reactivity index for the Ru atom[c]	−0.686
9	relative negative charge (RNCG)[d]	−0.627
8	total hybridization comp. of molecular dipole[b]	−0.619
4	average nucleophilic reactivity index for a C atom[e]	−0.536
19	moment of inertia B	−0.409
13	Ru $d_\pi \rightarrow L_\pi$ back-donation[f]	−0.290
21	HOMO−LUMO energy gap	0.131
17	max valency of a C atom	0.237
20	max bond order of a H atom	0.290
3	min nucleophilic reactivity index for a C atom	0.390
15	ZX shadow/ZX rectangle[g]	0.434
7	1-electron reactivity index for the Ru atom[h]	0.453
1	relative number of single bonds	0.620
5	average electrophilic reactivity index for a C atom	0.650
22	L→Ru σ-donation[i]	0.757
16	total dipole of the molecule (CHELPG)	0.900
12	steric exchange repulsion L-alkylidene[j]	0.976
14	Ru=CH_2 σ bond order[k]	1.101
10	Wiberg index for the Ru=CH_2 bond[l]	1.389

[a] Kier and Hall index (order 2) $= \sum_{i=1}^{N_{SB}} \prod_{k=1}^{3} (1/\delta_k^v)^{1/2}$ where $\delta_k^v = (Z_k^v - H_k)/(Z_k - Z_k^v - 1)$ is the valence connectivity for the k-th atom in the molecular graph. Z_k, the total number of electrons in the k-th atom; Z_k^v, the number of valence electrons in the k-th atom; H_k, the number of hydrogen atoms directly attached to the kth non-hydrogen atom.[73] [b] Lone pair contribution to molecular dipole vector. [c] Electrophilic Fukui reactivity index for the atom A, $E_A = \sum_{j\in A} c_{j\mathrm{LUMO}}^2/(\epsilon_{\mathrm{LUMO}} + 10)$. [d] RNCG $= \delta_{\max}/\sum_A \delta_A$, where $\delta_{\max}$ is the maximum atomic negative charge in the molecule and δ_A is a negative atomic charge in the molecule.[74] [e] Nucleophilic Fukui reactivity index for the atom A, $N_A = \sum_{i\in A} c_{i\mathrm{HOMO}}^2/(1 - \epsilon_{\mathrm{HOMO}})$. [f] The Ru $d_\pi \rightarrow L_\pi$ back donation is approximated as the difference between the NBO[52] 3d lone pair population on Ru in $LCl_2Ru{=}CH_2$ and the corresponding 3d lone pair population on Ru in $Cl_2Ru{=}CH_2$ (3.96 *e*). [g] By orientation of the molecule in the space along the axes of inertia (the *x* coordinate is along the main axis of inertia and so on), the areas of the shadows S1, S2, and S3 of the molecule as projected on the *XY*, *YZ*, and *XZ* planes are calculated.[75] The normalized shadow areas are calculated as the ratios S1/($X_{\max}Y_{\max}$), S2/($Y_{\max}Z_{\max}$), and S3/($X_{\max}Z_{\max}$), where $X_{\max}$, $Y_{\max}$, and $Z_{\max}$ are the maximum dimensions of the molecule along the corresponding axes. These descriptor indices therefore reflect the size (natural shadow indices) and geometrical shape (normalized shadow indices) of the molecule. [h] One-electron reactivity index for an atom A, $R_A = \sum_{i\in A}\sum_{j\in A} c_{i\mathrm{HOMO}} c_{j\mathrm{LUMO}}/(\epsilon_{\mathrm{LUMO}} - \epsilon_{\mathrm{HOMO}})$. [i] The L→Ru σ donation is approximated as the difference between the Ru $d_\pi \rightarrow L_\pi$ back donation and the combined charge[51] on the $Cl_2Ru{=}CH_2$ fragment. [j] Reference 55. [k] NBO[52] σ Ru=CH_2 bond order taken as BO $= 0.5(n_{\mathrm{occ}} - n^*_{\mathrm{occ}})$, where n_{occ} and n^*_{occ} are the occupancy numbers of the bonding and antibonding σ orbitals, respectively. [l] Reference 64. [m] The regression coefficients are obtained from autoscaled molecular descriptors (data matrix **X**).

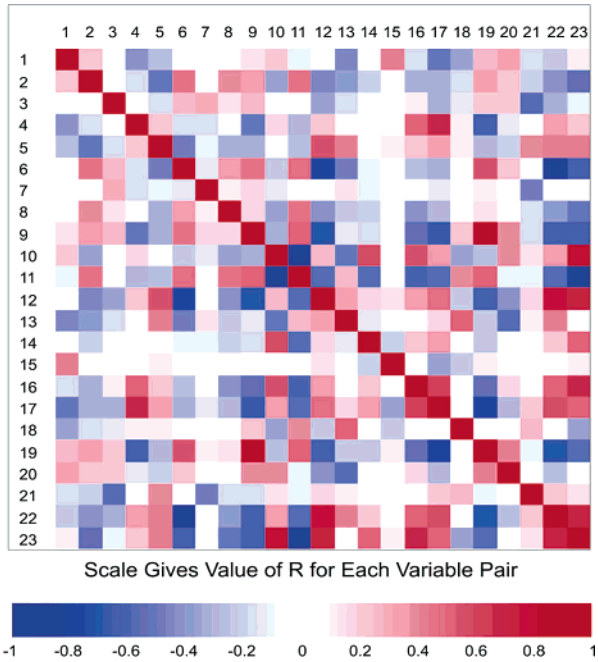


Figure 3. Correlation between the 22 independent variables (descriptors) and the response variable (productivity, variable no. 23, defined in Figure 1) retained in the final PLSR model. Numbering of descriptors: 1, relative number of single bonds; 2, Kier and Hall index (order 2); 3, minimum nucleophilic reactivity index for a C atom; 4, average nucleophilic reactivity index for a C atom; 5, average electrophilic reactivity index for a C atom; 6, electrophilic reactivity index for the Ru atom; 7, one-electron reactivity index for the Ru atom; 8, total hybridization component of the molecular dipole; 9, relative negative charge (RNCG); 10, Wiberg index for the Ru=CH_2 bond; 11, Ru=CH_2 bond distance; 12, steric exchange repulsion L−alkylidene; 13, Ru $d_\pi \rightarrow L_\pi$ back-donation; 14, Ru=CH_2 σ bond order; 15, ZX shadow/ZX rectangle; 16, total dipole of the molecule (CHELPG); 17, maximum valency of a C atom; 18, maximum bond order of a C atom; 19, moment of inertia B; 20, maximum bond order of a H atom; 21, HOMO−LUMO energy gap; 22, L→Ru σ-donation.

a higher effective oxidation state for the metal, closer to that of the metallacyclobutane intermediate (formally +4). A dative ligand suitable for olefin metathesis thus should be expected to promote the metal to a high effective oxidation state already at the entrance side of the reaction or, equivalently, to stabilize the oxidation state +4 of the metallacyclobutane intermediate (and the TS for formation thereof) with respect to the inactive 16-electron complex. If this line of reasoning is correct, we should expect to find ligands known for their low-oxidation state organometallic chemistry (e.g., CO) predominantly among the less active complexes (i.e., left-hand side of Figure 2), whereas ligands suitable for stabilization of higher oxidation states should be found among the more active catalysts. In fact, this is exactly what we observe. Tsipis et al.[24] recently pointed out that if such ligand stabilization of the high-oxidation state metallacyclobutane intermediate actually takes place, this stabilization should also be observed relative to the alkene π-complex (see Scheme 1) and not only relative to the 16-electron inactive complex. Whereas Tsipis et al.[24] did not observe stabilization of the metallacycle relative to the π-complex in their own calculations,[66] such preferential stabilization of the high-oxidation state intermediate can be discerned for the more active relative to the less active catalysts in other quantum chemical studies of the Ru-catalyzed olefin metathesis reaction mechanism where first- and second-generation Grubbs catalysts are compared.[19−23]

An effect of σ-donation from L is seen also on the charge of the alkylidene hydrogen atoms. The average charge on these hydrogen atoms correlates negatively with productivity ($R = -0.72$). This means that low positive partial charges (close to 0.2 *e*) on the alkylidene hydrogen atoms are found in the catalysts of high productivity, with strong L→Ru σ-donation, whereas significantly higher partial charges (close to 0.4 *e*) are obtained for the least active complexes; see Table 2. A corresponding effect of L→Ru σ-donation on the charge of the alkylidene carbon atom does not exist, however, suggesting that electron deficiency on the alkylidene carbon atom in catalysts

(66) In fact, inspection of the energies reported by Tsipis et al. reveals stabilization of transition states or metallacyclobutane intermediates compared to the alkene π-complexes for catalysts with NHC ligands compared to those of phosphines, albeit only very weakly so for their two intermediate models, $(PMe_3)_2Cl_2Ru{=}CHPh$ and $(PMe_3)((CH_2)_2(NH)_2C)$-$Cl_2Ru{=}CHPh$.

Table 2. Bond Orders (BO), Ligand−Metal Donation (L→Ru), π-Back-Donation (Ru $d_\pi \rightarrow L_\pi$), Charges (q), and Related Properties Calculated for Selected $LCl_2Ru{=}CH_2$ Complexes[a]

L[b]	net L → Ru	Ru $d_\pi \rightarrow L_\pi$	σ L → Ru	Ru=C BO	r(Ru=C)	avg Ru−Cl BO	avg q(Cl)[c]	q(C)[d]	avg q(H)[d]
CO (CO)	−0.03	0.35	0.32	1.536	1.835	0.567	−0.36	−0.40	0.20
P16 (PF_3)	0.02	0.17	0.18	1.539	1.833	0.581	−0.35	−0.39	0.20
P1 (PH_3)	0.24	0.08	0.32	1.604	1.824	0.543	−0.44	−0.42	0.17
P8 (PPh_3)	0.31	0.07	0.38	1.607	1.825	0.522	−0.44	−0.30	0.10
P7 (PCy_3)	0.32	0.06	0.38	1.641	1.819	0.501	−0.46	−0.27	0.12
A13 (IMes)	0.33	0.16	0.49	1.656	1.816	0.493	−0.47	−0.23	0.10
A14 (Cl_2IMes)	0.32	0.17	0.49	1.646	1.817	0.498	−0.45	−0.20	0.09
A12 (H_2IMes)	0.34	0.17	0.51	1.651	1.816	0.491	−0.47	−0.21	0.10
H1 (Pyridine)	0.25	0.09	0.34	1.644	1.822	0.517	−0.46	−0.34	0.13
N4 (Piperidine)	0.32	0.02	0.34	1.667	1.819	0.500	−0.47	−0.43	0.17
O4 (THF)	0.22	0.02	0.23	1.659	1.819	0.521	−0.46	−0.40	0.16

[a] Bond orders (BOs) are Wiberg bond indices.[64] Charges are given in multiples of the electron charge, and distances, in angstroms. [b] See Chart 2 for ligand structures. [c] Pertaining to the two chlorines of the $Cl_2Ru{=}CH_2$ fragment. [d] Pertaining to the alkylidene group.

with weak L→Ru σ-donation and strong Ru $d_\pi \rightarrow L_\pi$ back-donation is counteracted by electron flow from the alkylidene hydrogen atoms toward the carbon atom. For example, such charge compensation is seen for the CO ligand, for which the q_C(carbene) = −0.40 *e*, whereas q_C(carbene) > −0.30 *e* is obtained for much better donor ligands such as **A12** (H_2IMes) and **A13** (IMes). This charge compensation is also the reason for the seemingly counterintuitive positive correlation between productivity, L→Ru σ-donation, and the average electrophilic reactivity index for a C atom (descriptor no. 5, cf. Figure 3).

Whereas our absolute values for σ-donation and π-back-donation to the dative ligand (Ru $d_\pi \rightarrow L_\pi$) should be treated with care,[53] qualitative comparison between the different classes of ligand shows that the π-acceptor ability increases in the expected order for the classical ligands, e.g., ethers ≈ amines < imines ≈ phosphines < CO. However, whereas N-heterocyclic carbenes usually are referred to as virtually pure σ-donors,[65,67,68] our calculations suggest significant π-acceptor ability for these ligands. For example, π-back-donation for the NHC ligands is estimated to be ca. 0.15 *e* larger than that for typical amines, which have virtually no back-donation in these complexes and roughly half that of carbon monoxide (Table 2). With the exception of PF_3, the NHC ligands are also calculated to be better π-acceptors than the phosphines.[69] Some investigators have already noted indications of π-acidity for Arduengo-type carbenes,[67,70] and recently, evidence for non-negligible contribution from π back-donation to the metal−carbene bond has been reported for group 11 metals.[71,72] The present results suggest significant contribution from π-back-donation to the metal−NHC bond also in ruthenium-based catalysts for olefin metathesis.

In contrast to σ-donation, π-back-donation is negatively correlated with productivity, albeit only weakly, meaning that π-back-donation is not expected to have a large influence on catalytic activity (see Table 1). The weak correlation may seem puzzling given the significant span in values for π-back-donation obtained for the present ligands, from essentially zero (for, e.g., THF and amines) to 0.35 *e* (for CO). However, most of the dative ligands, and in particular those actually in use in the Grubbs catalysts, have calculated values for back-donation in the lower part of this range. π-Back-donation in these complexes is smaller and thus to a large extent masked by the excellent σ-donating abilities of their ligands. This is particularly the case for the N-heterocyclic carbenes which have the largest back-donation calculated for ligands actually in use in the Grubbs catalysts (Table 2). Most of these carbenes, including popular ligands such as IMes and H_2IMes, have larger calculated net L→Ru donation than the most basic phosphines.

Steric Effects

The influence of sterically demanding substituents, for example, mesityl substituents in the NHC ligands IMes and H_2IMes, has been the subject of recent controversy. Cavallo claimed that steric pressure exerted by such substituents on the alkylidene moiety in the precursor and the active 14-electron complex favors the metallacyclobutane intermediate and that this to a large extent explains the difference in activity between the first- and second-generation Grubbs catalysts.[21] In contrast, Adlhart and Chen maintained that the difference between the first- and second-generation catalysts "is to a great extent due to electronic effects, while steric differences, in particular between PCy_3 and H_2IMes, play a minor role".[22] To test explicitly the influence of steric pressure on the alkylidene moiety, we calculated the steric exchange interactions between the ligand, L, and the CH_2 moiety[54] using natural steric analysis[55] for all the active 14 e^- alkylidene complexes and included these energies as a descriptor (no. 12, Table 1) in the multivariate model.

It turns out that the steric exchange repulsion between the dative ligand, L, and the methylidene in the 14 e^- complex is strongly positively correlated with productivity. In other words, the most active catalysts have large calculated steric exchange repulsion between L and the CH_2-group[54] in the 14-electron active complexes. This relationship also showed up as a positive correlation between the angle L−Ru=CH_2 and the productivity in the initial PLSR model containing 122 molecular descriptors (these descriptors are listed in the Supporting Information). The

(67) Tafipolsky, M.; Scherer, W.; Ofele, K.; Artus, G.; Pedersen, B.; Herrmann, W. A.; McGrady, G. S. *J. Am. Chem. Soc.* **2002**, *124*, 5865.
(68) Despagnet-Ayoub, E.; Grubbs, R. H. *J. Am. Chem. Soc.* **2004**, *126*, 10198. Lee, M. T.; Hu, C. H. *Organometallics* **2004**, *23*, 976. Lai, C. L.; Guo, W. H.; Lee, M. T.; Hu, C. H. *J. Organomet. Chem.* **2005**, *690*, 5867.
(69) The back-donation to phosphines may be underestimated since the NBO method does not include 3d orbitals in the preselected valence space of phosphorus, as discussed in Frenking, G.; Frölich, N. *Chem. Rev.* **2000**, *100*, 717.
(70) Tulloch, A. A. D.; Danopoulos, A. A.; Kleinhenz, S.; Light, M. E.; Hursthouse, M. B.; Eastham, G. *Organometallics* **2001**, *20*, 2027.
(71) Hu, X. L.; Tang, Y. J.; Gantzel, P.; Meyer, K. *Organometallics* **2003**, *22*, 612. Nemcsok, D.; Wichmann, K.; Frenking, G. *Organometallics* **2004**, *23*, 3640. Hu, X. L.; Castro-Rodriguez, I.; Olsen, K.; Meyer, K. *Organometallics* **2004**, *23*, 755.
(72) Jacobsen, H. *J. Organomet. Chem.* **2005**, *690*, 6068.

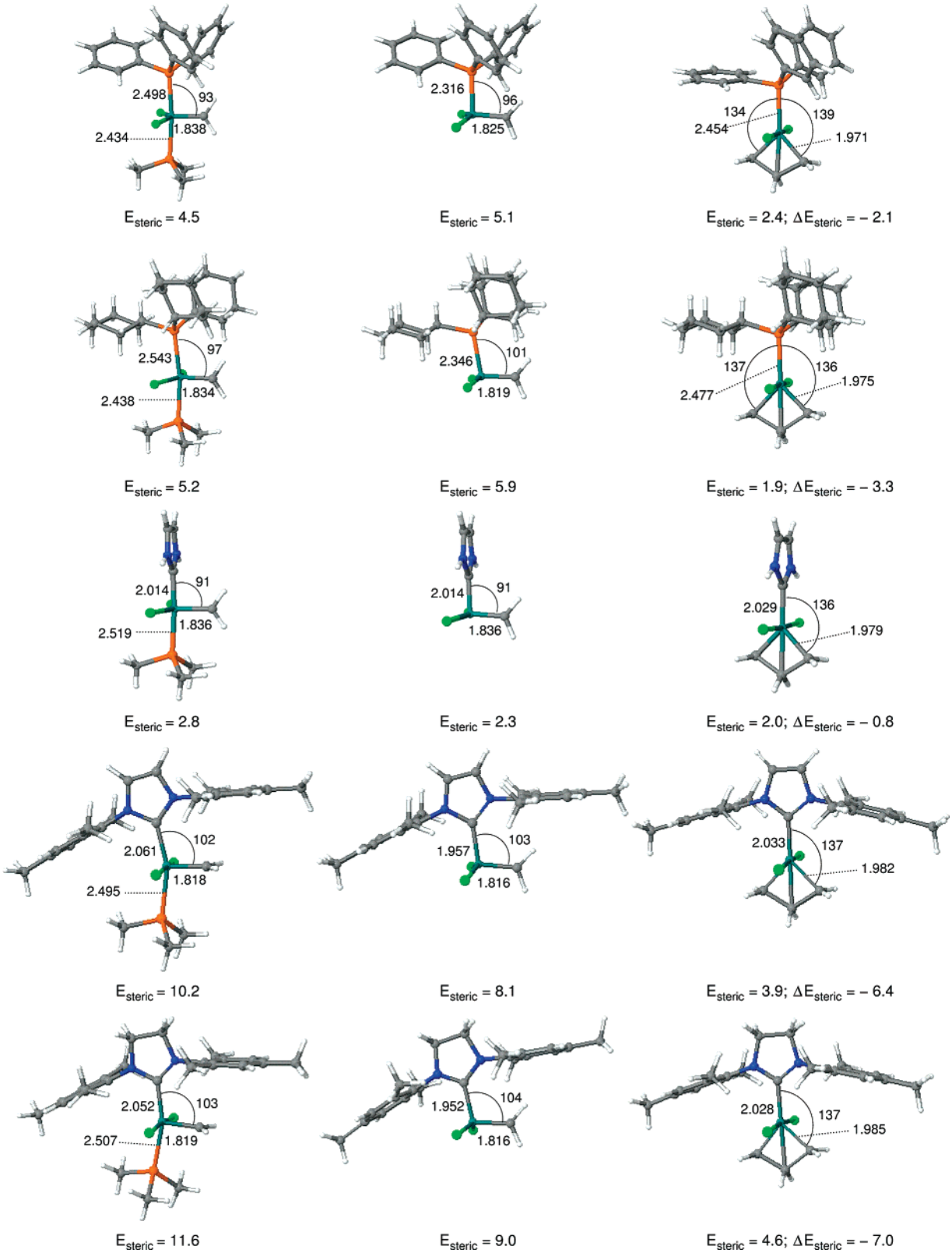


Figure 4. Selected bond distances [Å], angles [deg] and steric exchange repulsions, E_{steric} [kcal/mol], between the ligand, L, and the alkylidene moiety (metallacyclobutane moiety in the case of the metallacyclobutane intermediate)[54] in the 16-electron precursor complex (left), the active 14-electron complex (middle), and the metallacyclobutane intermediate (right) for the first (upper two rows) and second generation Grubbs catalysts (lower three rows). $\Delta E_{steric} = E_{steric}$(intermediate) $- E_{steric}$(precursor).

bulkier ligands give wider angles as a result of the increased steric repulsion toward the alkylidene group; see Figure 4. The correlation between steric repulsion and productivity suggests that the L−CH_2 steric repulsion is larger in the 16-electron precursor complex than in the metallacyclobutane intermediate and that increased steric pressure contributes to preferential stabilization of the intermediate and to increased activity, as proposed by Cavallo.[21] To verify this hypothesis we have

calculated the steric exchange repulsion L−CH_2 also for the 16-electron precursor and the metallacyclobutane intermediate for a series of catalysts, and these are given in Figure 4. Indeed, the steric exchange repulsion between the ligand, L, and the CH_2 moiety is seen to be much smaller in the metallacyclobutane intermediate than in both the precursor and the 14-electron active complex.[54] For example, with a triphenyl phosphine ligand (**P8**, PR_3, R = C_6H_5), the steric exchange repulsion, E_{steric}, in the intermediate is ca. 2 kcal/mol lower than that in the precursor ($\Delta E_{steric} = -2.1$ kcal/mol). Moreover, an increase in the steric repulsion calculated for the 14 e^- complex upon introducing a bulkier ligand, L, results in an increase in the steric repulsion also in the precursor, whereas the corresponding steric repulsion in the intermediate may increase only slightly and is even found to decrease in some cases. For example, exchanging R = C_6H_5 for R = C_6H_{11} in the phosphine ligand to give **P7** brings about an increase of ca. 0.9 (0.7) kcal/mol in steric exchange repulsion in the active (precursor) complex due to the bulkier cyclohexyl rings, whereas the repulsion actually decreases by ca. 0.5 kcal/mol in the metallacyclobutane intermediate. The ligand substitution thus results in a stabilization relative to the precursor, $\Delta\Delta E_{steric} = -1.2$ kcal/mol, and this stabilization is part of the reason for the higher catalytic activity observed[31] for catalysts of the tricyclohexylphosphine ligand.

Larger steric effects are, in turn, seen when exchanging phosphine for an NHC ligand, for example **A13** (IMes), which is accompanied by a relative stabilization of the intermediate, $\Delta\Delta E_{steric} = -3.1$ kcal/mol. For catalysts of the IMes and H_2IMes ligands, it should be noticed that the L−CH_2 steric repulsion is significantly larger in the precursor than in the 14-electron complex. Our calculations thus do not support the suggestion[21] that the olefin and phosphine free complexes of these ligands are particularly destabilized by steric strain which does "not promote phosphine dissociation" from the precursors of these catalysts.

Bulky N-bound substituents on the Arduengo carbene ensure that the ligand is oriented staggered to the chlorines in the equatorial plane and thus parallel to the Ru=CH_2 bond, as opposed to simple Arduengo-type carbenes such as **A2** (imidazol-2-ylidene), which are placed orthogonal to the Ru=CH_2 bond in order to form Cl−H hydrogen bonds.[21,24] The orthogonal orientation of the small carbenes gives little steric repulsion toward the alkylidene moiety, resulting in only small differences in steric exchange energy between the precursor and the metallacyclobutane intermediate, $\Delta E_{steric} = -0.8$ kcal/mol in the case of imidazol-2-ylidene, and probably also destabilizes the intermediate and the transition state region due the fact that, with this orientation of the NHC ligand, back-donation occurs from a Ru 4d orbital also participating in the Ru−C σ-bonds of the metallacycle.[24] Bulky N-bound substituents on Arduengo-type carbenes, on the other hand, ensure an orientation of the ligand parallel to the alkylidene bond and thus very specifically put steric pressure on the Ru=CH_2 moiety. In contrast, a small ligand such as CO has little steric repulsion toward the carbene moiety and only a small stabilizing effect on the metallacyclobutane intermediate, $\Delta E_{steric} = -1.5$ kcal/mol, thus to a large extent explaining why this and similar ligands are the least favorable for olefin metathesis reaction in our study (cf. Figure 4). Moreover, the specific steric pressure on the alkylidene moiety exerted by the NHC ligands, together with the excellent electronic properties of these ligands (vide supra), explains the higher catalytic activity experienced with Grubbs second generation than first generation catalysts.

There is a relative stabilization of the metallacyclobutane intermediate upon going from L = IMes to H_2IMes, $\Delta\Delta E_{steric} = -0.7$ kcal/mol. Thus, the higher catalytic activity generally seen for ruthenium second generation catalysts can to a large extent be explained by steric effects. Some of these steric effects are caused by the shorter Ru−NHC bond obtained for the saturated ligands (Figure 4). In other words, the steric effects in part originate from electronic differences between the saturated and unsaturated ligands. The better σ-donating and π-accepting abilities of the saturated ligands ensure a stronger and shorter Ru−NHC bond, which, in turn, causes closer contact and higher steric repulsion between the NHC ligand and the alkylidene moiety. This interplay of electronic and steric effects explains the higher catalytic activity observed for catalysts bearing H_2IMes, compared to those of IMes, although H_2IMes only is a marginally better σ-donor (as also pointed out in ref 61) and does not appear to be much bulkier than IMes.

Other Descriptors

Some descriptors correlate strongly with productivity but carry little or only indirect chemical information. The second-order Kier and Hall index, for example, correlates negatively with productivity, and this is mainly a result of the fact that heavy atoms contribute to large values for this descriptor, resulting in accidental correlation with more chemically informative descriptors. Thus, this descriptor correlates with the descriptors presented above providing the steric and electronic factors behind the difference in activity between Ru−phosphine and Ru−NHC complexes. High values for the Kier and Hall index are obtained for the phosphine complexes due to the presence of a third-row element in the ligand. Accidentally, this descriptor thus captures the difference between the first- and second-generation Grubbs catalysts. Another example of a descriptor that correlates accidentally with productivity is the relative negative charge (RNCG). Small relative negative charges are obtained for large molecules with many atoms carrying a negative charge. The negative correlation with productivity for this descriptor is thus simply a result of the fact that the ligands with the best donating and steric properties are large. Finally, several descriptors correlate only weakly with productivity but were retained in order to balance the multivariate model.

Prediction of New Catalysts

The structure−activity relationships presented above suggest which properties of the donor ligand, L, that should be promoted in order to increase activity for the Grubbs-type catalysts, LL′Cl_2Ru=CH_2. Moreover, our multivariate model represents an effective tool for cost-effective in silico testing of new donor ligands, placing us in an excellent position for generation and validation of ideas for the design of new dative ligands.

First, the Arduengo carbenes have been found to possess excellent steric as well as electronic properties for the olefin metathesis reaction (vide supra), and these ligands are dominating among the better catalysts in our comparison of productivities for a broad selection of complexes (Figure 2). In the present study we thus narrow our focus on potential new ligands to carbenes only. Our study already contains carbene ligands

that have not yet been synthesized. One such example is provided by the P-heterocyclic ligand (PHC, **C1**, Figure 2). While the present study was in progress, a stable PHC ligand and several of its rhodium complexes were reported.[76] Recent reports based on quantum chemical calculations suggest that this class of ligand should be competitive with the NHC ligands in applications as ligands for transition metal catalysts in general[72] and for ruthenium-based olefin metathesis catalysts in particular.[77] In contrast, the low productivity (−9.2 kcal/mol) predicted here for **C1** compared to the corresponding NHC ligand, **A2** (−3.2 kcal/mol), suggests a lower inherent olefin metathesis activity for ruthenium complexes of PHC ligands than that of NHC ligands. The low productivity of **C1** is partly due to the larger Ru $d_\pi \rightarrow L_\pi$ back-donation for **C1** (0.45 *e*) than for **A2** (0.14 *e*).

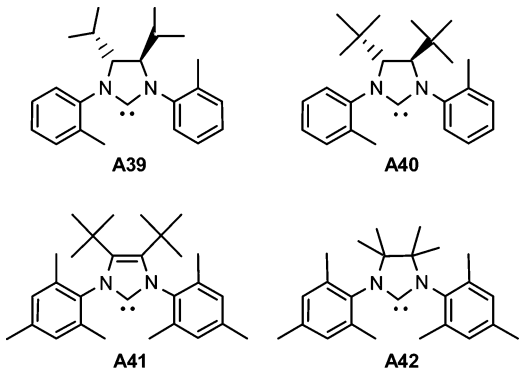
Turning now to the NHC ligands, the importance of both steric bulk and donation from L suggests that the nitrogen atoms should bear large and σ-donating substituents. Indeed, a catalyst based on an NHC ligand with *tert*-butyl-substituted nitrogen atoms (**A8**) has the highest predicted productivity (5.4 kcal/mol) of the complexes included in the multivariate model. However, as already noted above, the extreme steric requirements of the *tert*-butyl substituents weaken the Ru−L bond,[62] which most probably is the reason for the low stability observed for the catalyst of the closely related ligand **A7**.[63] The extent to which steric bulk can be increased by the subsituents on nitrogen is obviously limited if stability is to be maintained. In the present contribution we thus concentrate on substitution of the two backbone carbon atoms of the carbene ligand. The difference between H_2IMes and IMes ligands was found to be both electronic and steric in nature (vide supra), indicating the potential for further exploration of substituents at the backbone positions. Furthermore, three of the four catalysts with the highest predicted productivities in our multivariate model are substituted at the backbone carbon atoms; see Chart 2, Figure 2, and Table 3. Catalysts based on **A16** and **A20** have been synthesized and are very active for asymmetric olefin metathesis,[78] whereas **A17** and **A22**, shown in Chart 2, are new ligands originally included to ensure a wide variation span in our multivariate model. The four ligands **A16**, **A17**, **A20**, and **A22** all have identical substituents (*o*-tolyl) on the nitrogen atoms and thus offer insight into the effects of backbone substitution. Our calculations show that the two phenyl-substituted ligands (**A20** and **A22**) provide the highest steric exchange repulsion toward the alkylidene moiety of these four ligands, whereas the saturated alkyl substituted ligand **A16** is seen to be a better σ-donor than **A20**, **A22**, and **A17**. Further improvements in activity might thus be anticipated by incorporation of excellent σ-donor substituents with large steric requirements, for example, iso-propyl and *tert*-butyl groups as in **A39**−**A41**, or four methyl groups as in **A42** (see Chart 3). Except for the special case of **A8**,[62,63] complexes of **A39**−**A42** give the highest productivities of all complexes in the present investigation, suggesting that these ligands, or other carbene ligands with the backbone carbon atoms substituted by alkyl groups or other sterically demanding and σ-donating substituents, could provide catalysts that are more active and functional group tolerant than contemporary olefin metathesis catalysts, provided that these catalysts can be synthesized and are sufficiently stable.[79]

Table 3. Predicted and Explicitly Calculated Productivities of Existing and New Catalysts, $LL'Cl_2Ru{=}CH_2$[a]

ligand[b]	predicted productivity	calculated productivity
	Existing Catalysts	
P8 (PPh_3)	−10.7	−11.3
P7 (PCy_3)	−3.5	−2.9
A12 (H_2IMes)	2.4	2.6
A13 (IMes)	2.3	1.8
A14 (Cl_2IMes)	1.1	1.2
A16	2.5	2.8
A20	4.1	3.7
	Predicted New Catalysts	
A17	3.8	
A22	2.1	
A39	5.9	4.5
A40	5.9	4.6
A41	5.1	4.5
A42	4.4	4.6

[a] Productivities in kcal/mol; see Figure 1 for definition. [b] See Chart 2 for ligand structures.

Chart 3. New NHC Ligands, L, Predicted To Provide Highly Active Catalysts, $LL'Cl_2Ru{=}CH_2$, for the Olefin Metathesis Reaction; See Table 3



Conclusions

We have shown that a multivariate QSAR model for which both the independent and dependent (response) variables are derived from DFT calculations is able to capture the experimentally recorded order of activity for the Grubbs ruthenium catalysts for olefin metathesis. The QSAR model correlates the properties of a large number of active 14-electron complexes with a calculated measure of activity for a limited number of optimally selected, statistically representative complexes, with high accuracy, and is seen to be a cost-effective approach in the absence of a large set of comparable experimental activities. The accuracy and applicability of the model is to a large extent due to the use of highly specific geometric and electronic molecular descriptors which give unprecedented insight into the

(73) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; J. Wiley & Sons: New York, 1986.
(74) Stanton, D. T.; Jurs, P. C. *Anal. Chem.* **1990**, *62*, 2323.
(75) Rohrbaugh, R. H.; Jurs, P. C. *Anal. Chim. Acta* **1987**, *199*, 99.
(76) Martin, D.; Baceiredo, A.; Gornitzka, H.; Schoeller, W. W.; Bertrand, G. *Angew. Chem., Int. Ed.* **2005**, *44*, 1700.
(77) Schoeller, W. W.; Schroeder, D.; Rozhenko, A. B. *J. Organomet. Chem.* **2005**, *690*, 6079.
(78) Seiders, T. J.; Ward, D. W.; Grubbs, R. H. *Org. Lett.* **2001**, *3*, 3225.

(79) The calculated Ru−L bond dissociation enthalpies of ligands **A39**−**A42** with the fragment $Cl_2Ru{=}CH_2$ are similar to those of NHC ligands such as IMes and H_2IMes, suggesting that the catalyst complexes of **A39**−**A42** should have stabilities comparable to those of contemporary Grubbs catalysts for olefin metathesis. The synthesis of the predicted catalysts based on **A39**−**A42** is currently being pursued in our laboratories.

factors governing catalytic activity and establish a direct connection between activity and chemically meaningful donor ligand properties.

The ligands that most efficiently promote catalytic activity are those that stabilize the high-oxidation state (+4) metallacyclobutane intermediate, and the accompanying transition states for ring-closure and -opening, relative to the ruthenium−carbene structures dominating the rest of the reaction pathway. Stabilization of the intermediate is ensured through ligand-to-metal σ-donation, which thus correlates strongly with catalyst activity, whereas metal-to-ligand π-back-donation contributes to lowering the activity. The steric repulsion between the ligand and the alkylidene moiety also correlates strongly with catalytic activity, and the presence of a bulky dative ligand drives the reaction toward the less sterically congested metallacyclobutane species.

In addition to reproducing the known order of activity between existing catalysts and explaining the excellent performance of contemporary ruthenium-based catalysts, the multivariate model and its high-level descriptors provide practical handles for further catalyst development. The potential for development is exemplified by the suggestion of several new donor ligands predicted to give even more active and functional group tolerant catalysts for the olefin metathesis reaction.[79] Finally, the strategy used in the current work holds great promise for broader screenings of olefin metathesis catalysts as well as for development of homogeneous transition metal catalysts in general.

Acknowledgment. The Norwegian Research Council is gratefully acknowledged for financial support through the KOSK program (Project No. 160072/V30) as well as for CPU resources granted through the NOTUR supercomputing program.

Supporting Information Available: Complete ref 37, a list of all the molecular descriptors, the numerical values of key descriptors, predicted and explicitly calculated productivities, dendrogram of the cluster analysis, and additional results from and information about the PCA of the initial set of complexes (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA060832I