

ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



www.rsc.org/chemcomm

COMMUNICATION

Development of a web-based platform for studying lithiation reactions *in silico*.

Cite this: DOI: 10.1039/x0xx00000x

Mikhail A. Kabeshov,^{a,*} Eric Sliwinski,^a Daniel E. Fitzpatrick,^a Biagia Musio,^a James A. Newby,^a Wayne D. W. Blaylock,^b Steven V. Ley^aReceived 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

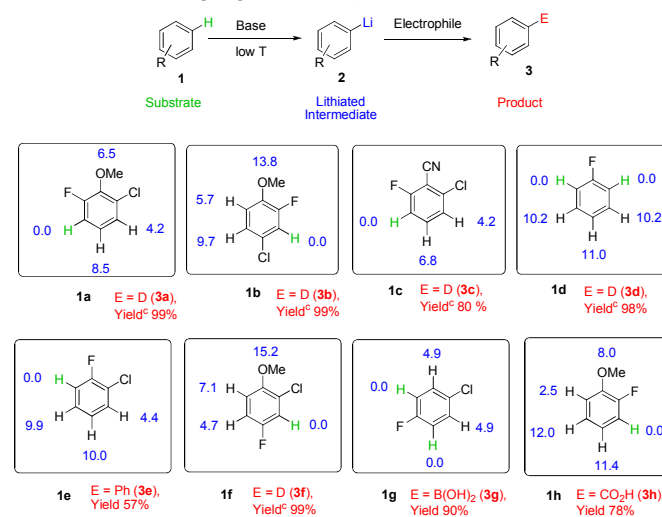
www.rsc.org/

A novel integrated web-based system which can compute, visualise and store systematised key parameters of a reaction has been developed from open-source components. As a proof of concept, it has been used to rationalise and predict the regioselectivity of lithiation reactions as well as relative reactivity of substrates. Excellent agreement between the *in silico* analysis and experimental data was obtained.

Understanding reactivity is fundamental to comprehend complex concepts of chemistry.¹ Therefore, building rational, widely applicable systems to rationalise, tune and predict outcomes of chemical reactions remains at the forefront of discovery.² Rapid development of IT and computer science technologies makes it possible to create new-age intelligent libraries and databases which can process, evaluate and generate new data. Increased availability and reliability of open-source software allows the development of highly versatile multi-tasking systems that are easily adjustable for a particular use. In the future we believe our laboratories will be increasingly controlled by computer systems and machine based enabling technologies to improve safety and efficiency.³

Directed lithiation is an example of a strategically important process ultimately leading to the formation of new bonds and hence to advanced molecular structures.⁴ Recent advances in flow and flash chemistry have provided new opportunities for the generation of reactive anions via deprotonation with organolithium reagents and their subsequent quenching with electrophiles under various tuneable conditions with the machine control.⁵ At the same time, comprehensive physical organic and computational studies have revealed many new features of the lithiation process.^{4b,6} The next advance in these methods would be to predict the outcome of a reaction prior to experiment quickly and reliably, thereby reducing labour and material costs. Such an approach could be realised using an integrated system – a science gateway – which would enable researchers to submit structures of interest for the computational analysis and subsequently retrieve the results which could be also

stored in the web-accessible database.⁷ By implementing this workflow on a webserver, users should be able to access, validate and visualise proposed reaction intermediates when needed from any location in the world. Here we report the first generation of a webserver-based automatic generic algorithm, organic chemistry science gateway, employing Density Functional Theory (DFT) level calculations to rationalise and predict the relative reactivity of aromatic compounds during lithiation reactions and the regioselectivity of subsequent quenching by electrophiles.

Scheme 1. Predicting regioselectivity of lithiation reactions.^{a,b}

^aProtons marked green are the ones that react during the lithiation step under specified conditions; Corrected electronic energies (computed at B3LYP/6-31g(d,p) + SMD solvation; the energy of the most stable regioisomer set at 0.0 kcal mol⁻¹) of all lithiated intermediates **2** are shown as blue; the group introduced during the electrophilic quenching and the overall yield of the 2-step process are shown as red.^b Lithiation conditions: **1a**, **1f**: n-BuLi/TMEDA (2 eqv), -78 °C, 2 h; **1b**, **1d**: n-BuLi/TMEDA (2 eqv), -78 °C, 1 h; **1c**: LDA (2 eqv), -78 °C, 2 h; **1e**: LiTMP (4 eqv), -78 °C, 18 h;⁸ **1g**: LiTMP (1.1 eqv), -75 °C, 5 h, with in situ quenching;⁹ **1h**: n-BuLi/PMDTA (1 eqv), -78 °C, 2 h.¹⁰ ^cDetermined by ¹H NMR.

For the DFT calculations, we employed NWChem software package which is an open-source DFT and *ab initio* quantum-chemistry program. It is easily scalable and designed to treat large scientific computational problems efficiently working on modern supercomputer clusters.³¹ Initially, the relative stability of lithiated intermediates was computed for a series of aromatic compounds, assuming³² that the aryl lithium intermediates exist mostly as monomers in THF solution and with the approximation that there are no specific kinetic effects associated with organolithium solvated clusters. To model these systems, every hydrogen atom in a molecule **1** was sequentially substituted by a lithium atom, and then the corrected electronic energies of regioisomers **2** were computed at B₃LYP/6-31g(d,p) level³³ in combination with an implicit solvation model(SMD) (Scheme 1).³⁴

The lithiated intermediate **2** with the lowest computed energy is predicted to be the one formed under the reaction conditions and to give the observed regioisomers of the product **3**. The ability of the proposed model to predict the regiochemistry of directed lithiation reactions followed by quenching with electrophiles has been successfully tested on a number of substituted arene derivatives (Scheme 1).

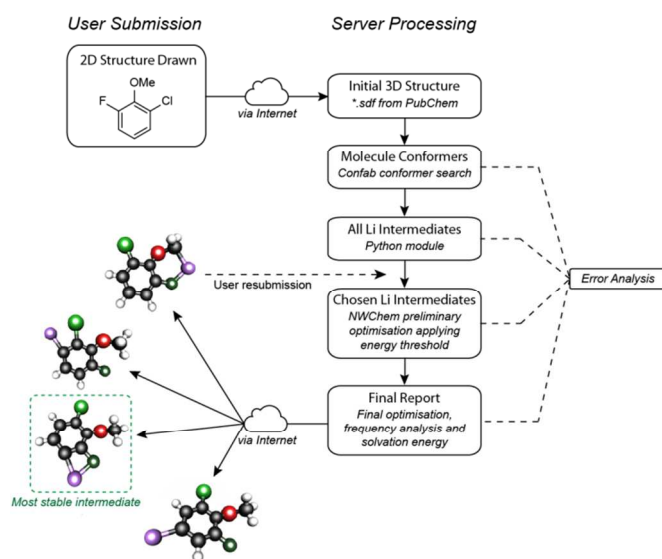


Figure 1. The automated algorithm for predicting lithiation reactions outcome.

A predictive algorithm that can operate *in silico* with extensive interaction was designed for the automated implementation of the proposed computational model (Figure 1).³⁵ Through a web-interface, users draw a 2D structure which is then processed³⁶ to generate 3D structures of the most stable conformers.³⁷ Subsequently, all possible lithium intermediates were built for each conformer. Quick geometry optimisation³⁸ was used to discard a large number of high energy intermediates leaving only the most plausible ones. Full geometry optimisation of the selected most stable intermediates was then performed and their electronic energies³⁹ were listed (Figure 1). It was not unusual for both the geometry optimization and the electronic structure DFT calculations to fail, or to converge to a local maximum

on the potential energy surface. In order to make these computations fully automatic and robust, a number of error checks on the calculations, with restarts using different input options and initial guesses were added.

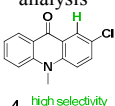
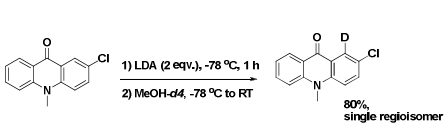
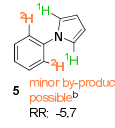
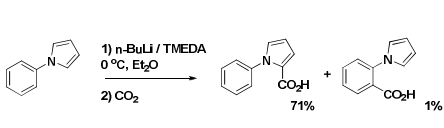
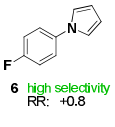
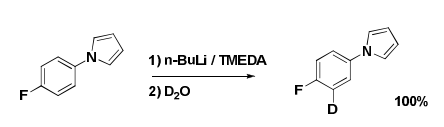
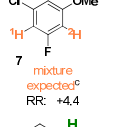
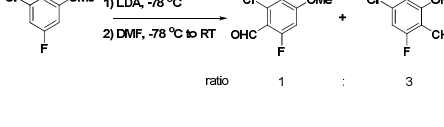
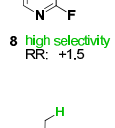
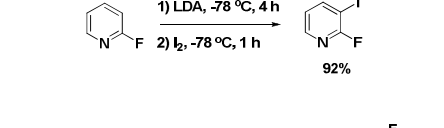
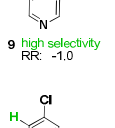
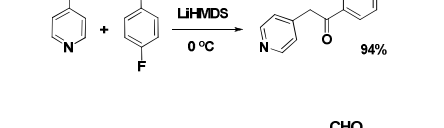
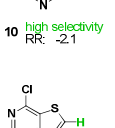
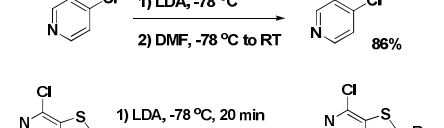
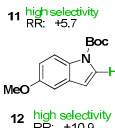
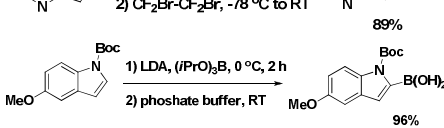
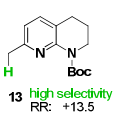
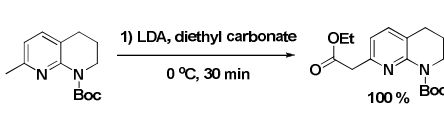
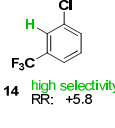
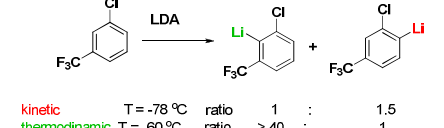
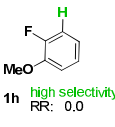
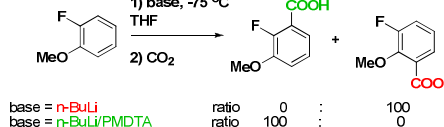


Each likely intermediate generated is then rendered as a 3D interactive structure applying HTML5 technologies with the help of open source three.js library.²⁰ A JavaScript file, consisting of an array of atoms²¹ and an array of bonds, is generated for each structure. This JavaScript file is then passed to a client-side HTML viewer where each atom is rendered using WebGL at its location with the corresponding atomic radius²² and CPK colouring.²³ Non-general elements²⁴ are subsequently labelled and the bonds rendered. The resulting 3D scene is then fully interactive, with the possibility for the user to zoom, pan or rotate the structure of the intermediate, hide the hydrogen atoms or the bonds, modify the structure representation (balls & sticks, sticks or compact), measure interatomic distances, angles and dihedral angles, edit the 3D structure by moving single atoms and export the modified coordinates and export the 3D structure as JSON or STL models. If needed, the user-modified structure can be resubmitted for computational analysis. Energies of automatically generated and user-modified (if applicable) intermediate structures for each analysed molecule together with its relative reactivity are also presented.

This algorithm can be used to rationalize (compared with experimental results reported in the literature; Entries 2 – 12, Table 1) as well as predict (Entry 1, Table 1) behaviour of other aromatic and heteroaromatic substrates. When the most stable intermediate **2** is significantly lower in energy³⁵ than other possible ones for a substrate, high regioselectivity of the lithiation reaction is anticipated (Entries 1, 3, 5 – 12; Table 1). In other cases, formation of more than one regioisomer is expected (Entries 2 and 4; Table 1). In all tested cases, the system is invariant of an electrophile used to quench the incipient anions (Table 1).

Corrected electronic energies for the formation of lithiated intermediates can also be compared between different substrates. Among other factors,²⁶ it has a great significance in defining relative reactivity of various compounds in the lithiation reaction and thus can serve as foundation in building a reactivity scale for molecules in lithiation reactions where a *reactivity ranking* (RR, Table 1) would be assigned to each of them. At this stage, we approximate RR as the corrected electronic energy of the lithiation relative to a chosen standard.²⁷ Substrates characterised with relatively high RR values were classified as “reactive” and could be successfully lithiated with LDA (Entries 1, 4, 5, 8 - 11; Table 1) whereas stronger n-BuLi/TMEDA was used as lithiating agent in cases with lower reactivity rankings (Entries 2, 3, 12; Table 1).

As it can be anticipated knowing the approximations used, the model in its current stage has a number of limitations. Thus, it can describe systems under thermodynamic control (Entry 11; Table 1) and where there are no specific kinetic effects derived from the state of oligomerisation of a lithiating reagent (Entry 12; Table 1). Also, as the spin-restricted DFT methods are currently used within the workflow, species with multiconfigurational character cannot be adequately described.²⁸ Work on expanding the scope of the method as well as increasing its accuracy is ongoing.

Table 1. Selected examples of lithiation reaction analysis and prediction.^a

Entry	<i>In silico</i> analysis	Experimental results
1		
2 ²⁹		
3 ³⁰		
4 ³¹		
5 ³²		
6 ³³		
7 ³⁴		
8 ³⁵		
9 ³⁶		
10 ³⁷		
11 ²⁶		
12 ³⁸		

^a Removal of the proton marked green leads to the formation of the most stable lithiated intermediate **2**; “high selectivity” refers to cases where the most stable intermediate is at least 2.0 kcal mol⁻¹ more stable than any other (corrected electronic energies are compared); “RR” stands for “reactivity ranking” and is the corrected electronic energy of the lithiation relative to the lithiation of **1h** chosen as a standard. ^b ¹H-intermediate is more stable than ²H by 1.3 kcal mol⁻¹. ^c Energy difference between intermediates ¹H and ²H is 0.2 kcal mol⁻¹.

Conclusions

An automated algorithm which can be used to rationalise and predict the regioselectivity of lithiation reactions as well as the relative reactivity of substrates has been developed and implemented as an open-source computational and visualisation system. The new algorithm does not require specialist knowledge or experience in computational chemistry and is designed to assist experimentalists by improving efficiency and productivity as well as deepening understanding of molecular reactivity.¹¹ Future work seeks to expand the scope and accuracy of the *in silico* model. We are also planning to describe other types of directed metallation in similar fashion.

Notes and references

^a Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW; E-mail: mk787@cam.ac.uk; <http://leygroup.ch.cam.ac.uk>.

^b Dow Chemical Company, Midland, Michigan 48674, USA

† The authors would like to thank Prof. Jonathan Goodman, Dr Peter Murray-Rust, Dr. Richard Turner, Dr Serghei Chercheja and Dr Ion Chiorescu for helpful discussions and suggestions. Also, the authors are thankful to the EPSRC (Award Nos. EP/K009494/1 and EP/K039520/1) and Dow for financial support.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/c000000x/

¹ IUPAC.Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford, 1997.

² a) M. Kayala, C.-A. Azencott, J. H. Chen, and P. Baldi, *J. Chem. Inform. Modeling*, 2011, **51**, 2209; b) W. Warr, *Mol. Inform.*, 2014, **33**, 469 and references therein; c) I. M. Socorro, K. Taylor and J. M. Goodman, *Org. Lett.*, 2005, **7**, 3541; d) H. S. Rzepa, *J. Cheminformatics*, 2013, **3**, 46

³ a) I. R. Baxendale, and S. V. Ley, Ernst Schering Foundation Symposium Proceedings 06.3, 2007, **3**, 151; b) R. J. Ingham, C. Battilocchio, J. M. Hawkins and S. V. Ley, *Beilstein J. Org. Chem.*, 2014, **10**, 641; c) S. V. Ley, R. J. Ingham, M. O'Brien and D. L. Browne, *Beilstein J. Org. Chem.*, 2013, **9**, 1051.

⁴ a) J. Board, J. L. Cosman, T. Rantanen, S. P. Singh and V. Snieckus, *Platinum Metals Rev.*, 2013, **57**, 234; b) M. C. Whisler, S. MacNeil, V. Snieckus and P. Beak, *Angew. Chem. Int. Ed.*, 2004, **43**, 2206; c) G. Rouquet, D. C. Blakemore and S. V. Ley, *Chem. Comm.*, 2014, **50**, 8908.

⁵ a) J.-I. Yoshida, “Flash Chemistry: Fast Organic Synthesis in Microsystems”, Wiley, 2008, ISBN-10: 0470035862; and references therein; b) A. Nagaki, D. Ichinari and J.-I. Yoshida, *J. Am. Chem. Soc.*, 2014, **136**, 12245. c) J. A. Newby, L. Huck, D. W. Blaylock, P.

- M. Witt, S. V. Ley and D. L. Browne, *Chem. Eur. J.*, 2014, **20**, 263.
- d) J. A. Newby, D. W. Blaylock, P. M. Witt, J. C. Pastre, M. K. Zacharova, S. V. Ley and D. L. Browne, *Org. Proc. Res. Dev.*, 2014, 1211. e) J. A. Newby, D. W. Blaylock, P. M. Witt, R. M. Turner, P. L. Heider, B. H. Harji, D. L. Browne and S. V. Ley, *Org. Proc. Res. Dev.*, 2014, 1221.
- ⁶a) D. B. Collum, A. J. McNeil and A. Ramirez, *Angew. Chem. Int. Ed.*, 2007, **46**, 3002; b) J. M. Saa, J. Morey, A. Frontera and P. M. Deya, *J. Am. Chem. Soc.*, 1995, **117**, 1105. c) T. Balle, M. Begtrup, J. W. Jaroszewski, T. Liljefors and P.-O. Norrby, *Org. Biomol. Chem.*, 2006, **4**, 1261. d) D. Tilly, J. Magolan and J. Mortier, *Chem. Eur. J.*, 2012, **18**, 3804. For DFT studies on similar systems involving Ru(II) complexes, see: I. Ozdemir, S. Demir, B. Cuetinkaya, C. Gourlaouen, F. Maseras, C. Bruneau and P. H. Dixneuf, *J. Am. Chem. Soc.*, 2008, **130**, 1156.
- ⁷ For other examples of science gateways, see a) J. Krüger, R. Grunzke, S. Gesing, S. Breuers, A. Brinkmann, L. De La Garza, O. Kohlbacher, M. Kruse, W. E. Nagel, L. Packschies, R. Müller-Pfefferkorn, P. Schäfer, C. Schärfe, T. Steinke, T. Schlemmer, K. D. Warzecha, A. Zink and S. Herres-Pawlis, *J. Chem. Theory Comp.*, 2014, **10**, 2232 and references therein; b) R. Dooley, K. Milfeld, C. Guiang, S. Pamidighantam and G. Allen, *J. Grid Comput.*, 2006, **4**, 195.
- ⁸ T. Truong, M. Mesgar, K. K. A. Le and O. Dauguils, *J. Am. Chem. Soc.*, 2014, **136**, 8568.
- ⁹ M. Cameron, B. S. Foster, J. E. Lynch, Y.-J. Shi and U.-H. Dolling, *Org. Proc. Res. Dev.*, 2006, **10**, 398.
- ¹⁰ M. A. Pena, J. P. Sestelo and L. A. Sarandeses, *J. Org. Chem.*, 2007, **72**, 1271.
- ¹¹ M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus and W. A. de Jong, *Comput. Phys. Commun.*, 2010, **181**, 1477. For usage, see: a) S. Kolakkandy, S. Pratihari, A. J. A. Aquino, H. Wang and W. L. Hase, *J. Phys. Chem. A*, 2014, **118**, 950; b) E. Yoo, D. B. Salunke, D. Sil, X. Guo, Alex C. D. Salyer, A. R. Hermanson, M. Kumar, S. S. Malladi, R. Balakrishna, W. H. Thompson, H. Tanji, U. Ohto, T. Shimizu and S. A. David, *J. Med. Chem.*, 2014, **57**, 7955.
- ¹² For related NMR studies, see: J. C. Riggs, A. Ramirez, M. E. Creameens, C. G. Bashore, J. Candler, M. C. Wirtz, J. W. Coe and D. B. Collum, *J. Am. Chem. Soc.*, 2008, **130**, 3406. For related computational studies and use of the same approximation when organolithium oligomers were detected, see 6c.
- ¹³ A. D. Becke, *J. Chem. Phys.*, 1997, **107**, 8554. For benchmarking studies and analysis of the explicit solvation effect, see ref. 6c.
- ¹⁴ For SMD solvation model, see: A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378.
- ¹⁵ Access to the computational system can be granted on request from the authors; for more details please see: <http://www.leygroup.ch.cam.ac.uk/research/science-gateway>
- ¹⁶ Confab, a systematic conformational search tool, was used to generate most stable conformers. For details, please see: N. M. O'Boyle, T. Vandermeersch, C. J. Flynn, A. R. Maguire and G. R. Hutchison, *J. Chemoinform.*, 2011, **3**, 8.
- ¹⁷ The energy cut-off of 10 kcal mol⁻¹ was used
- ¹⁸ Maximum number of steps during the geometry optimisation is set to 5
- ¹⁹ Electronic energies are after applying zero-point implicit solvation corrections
- ²⁰ three.js_rev.69 is freely available to download from the web: <https://github.com/mrdoob/three.js/>
- ²¹ Each atom object contains element and its cartesian coordinates
- ²² For more details, see: http://en.wikipedia.org/wiki/CPK_coloring
- ²³ The SMD intrinsic atomic Coulomb radii were used as implemented in NWChem software. For more details, see: http://www.nwchem-sw.org/index.php/Release65:SMD_Model
- ²⁴ All other than other than C, H, O, N
- ²⁵ Threshold of 2.0 kcal mol⁻¹ was used
- ²⁶ A. C. Hoepker, L. Gupta, Y. Ma, M. F. Faggini and D. B. Collum, *J. Am. Chem. Soc.*, 2011, **133**, 7135.
- ²⁷ *o*-Fluoroanisole **1h** was chosen as the standard, and the corrected electronic energy of its lithiation was set at zero.
- ²⁸ For example, some types of annulenes cannot be described as closed shell species. See: W. C. McKee, J. I. Wu, H. S. Rzepa and P. von R. Schleyer, *Org. Lett.*, 2013, **15**, 3432 and references therein. Currently, the system is limited not to describe charged systems and transition metal complexes
- ²⁹ F. Faigl and M. Schlosser, *Tetrahedron*, 1993, **49**, 10271.
- ³⁰ F. Faigl, K. Fogassy, Z. Szanto, A. Lopata and L. Toeke, *Tetrahedron*, 1998, **54**, 4367.
- ³¹ L. Gupta, A. C. Hoepker, K. J. Singh and D. B. Collum, *J. Org. Chem.* 2009, **74**, 2231.
- ³² M. Zhu, Z. Qiu, G. P. Hiel and S. Sieburth, *J. Org. Chem.*, 2002, **67**, 3487.
- ³³ F. Venturoni, N. Nikbin, S. V. Ley and I. R. Baxendale, *Org. Biomol. Chem.*, 2010, **8**, 1798.
- ³⁴ M. Kusaka, Patent EP1348706 A1, 2003.
- ³⁵ B.-S. Yang, Y.-Z. Kim, T.-K. Park, S.-H. Woo, H.-S. Lee, S.-Y. Kim, J.-U. Cho and H.-B. Lee, Patent: US2013/72482 A1, 2013.
- ³⁶ M. Mizuta, K. Seio, K. Miyata and M. Sekine, *J. Org. Chem.*, 2007, **72**, 5046.
- ³⁷ W. H. Miller, P. J. Manley, R. D. Cousins, K. F. Erhard, D. A. Hearding, C. Kwon, S. T. Ross, J. M. Samanen, D. T. Takata, I. N. Uzinskas, C. C. K. Yuan, R. C. Haltiwanger, C. J. Gress, M. W. Lark, S.-M. Hwang, I. E. James, D. J. Rieman, R. N. Willette, T.-L. Yue, L. M. Azzarano, K. L. Salyers, B. R. Smith, K. W. Ward, K. O. Johanson and W. F. Huffman, *Bioorg. Med. Chem. Letters*, 2003, **13**(8), 1483.
- ³⁸ G. Katsoulos, S. Takagishi and M. Schlosser, *Synlett*, 1991, 731.