

VIP Molecular Diversity Very Important Paper

Synthesis of Complex and Diverse Compounds through Ring Distortion of Abietic Acid**

Ryan J. Rafferty, Robert W. Hicklin, Katherine A. Maloof, and Paul J. Hergenrother*

Abstract: Many compound screening collections are populated by members that possess a low degree of structural complexity. In an effort to generate compounds that are both complex and diverse, we have developed a strategy that uses natural products as a starting point for complex molecule synthesis. Herein we apply this complexity-to-diversity approach to abietic acid, an abundant natural product used commercially in paints, varnishes, and lacquers. From abietic acid we synthesize a collection of complex (as assessed by fraction of sp³-hybridized carbons and number of stereogenic centers) and diverse (as assessed by Tanimoto analysis) small molecules. The 84 compounds constructed herein, and those created through similar efforts, should find utility in a variety of biological screens.

High-throughput screening (HTS) is widely used in the identification of new drug leads. Prior to the advent of HTS, compound collections were typically enriched in natural products and/or extracts from natural sources. Advances in robotics and parallel/combinatorial synthesis in the early 1990s led to a surge in the use of HTS, the results of which have been significant: for sorafenib, maraviroc, rivaroxaban, imatinib, lapatinib, ambrisentan, and several other FDA-approved drugs, the initial leads were identified through a high-throughput screen. [1-3]

HTS collections are generally populated by highly planar compounds, molecules with a significant percentage of sp²-hybridized carbon atoms and few stereogenic centers. [4-8] These small, flat compounds are useful against some biological targets; for example, three of the drugs mentioned above (sorafenib, imatinib, lapatinib) inhibit kinases, a class of targets well-suited for modulation by compounds that mimic the planar base of ATP. There are other cellular targets (e.g., PARP-1^[9]) that also appear to be effectively bound by planar compounds with little or no stereochemical complexity. However, it is likely that such planar compounds will be unsuitable for the modulation of certain challenging biological targets, including protein–protein interactions^[10] and

transcription factors.^[11] In addition, drugs for some indications (such as bacterial infections) are almost always complex molecules, possessing physiochemical properties very different from those normally found in HTS collections.^[12–14]

Natural products are (typically) complex molecules that have enjoyed notable success in drug discovery. [15] Such compounds possess considerable three-dimensionality, display a diversity of chemical functionality, and can modulate biological systems in ways that simpler compounds often cannot. For example, FK506 (Tacrolimus) binds simultaneously to the proteins FKBP12 and calcineurin, [16] vinblastine binds tubulin and inhibits its assembly into microtubules,[17] and ET-743 (Yondelis) selectively and covalently alkylates specific sequences within the minor groove of DNA.[18] Unfortunately, the identification of new natural products is challenging, [19,20] thus the source of complex compounds for future drug discovery efforts is uncertain. To fill this gap, several creative strategies for the generation of such compounds have been devised, including diversity-oriented synthesis, [21-27] biology-oriented synthesis, [28] synthesis of naturalproduct-inspired scaffolds, [29-32] synthesis of chiral, conformationally constrained oligomers, [33,34] and synthesis of compounds to probe biological/chemical space.^[35]

With the goal of creating structurally diverse compound screening libraries populated by complex molecules, we recently reported a general strategy for the synthesis of complex and diverse compound collections from readily available natural products.^[36] This approach, termed complexity-to-diversity (CtD), was used to prepare over 160 compounds from gibberellic acid, adrenosterone, and quinine.^[36]

As changes in the molecular scaffold lead to the most dramatic changes in overall topology, [37] CtD is focused on achieving scaffold diversity through the manipulation of core ring systems. Reaction sequences are designed to rapidly and dramatically alter ring systems through the use of ring distortion reactions. These are chemical transformations that directly alter the composition of rings in a molecule and include reactions that change ring size (ring expansion and ring contraction), form or break rings (ring fusion and ring cleavage), manipulate multiple rings (ring rearrangement), and aromatize rings. The strategic application of ring distortion reactions to complex natural products provides an efficient approach for the generation of highly diverse collections of complex small molecules in few synthetic steps. Here, we apply CtD to the diterpene abietic acid (AA, Scheme 1) and construct 84 complex compounds from this natural product.

AA is the major component of pine tree rosin and has been used for centuries in varnishes, lacquers, and ship

^[**] We are grateful to the Office of Naval Research (N00014-09-1-0249) and the University of Illinois at Urbana-Champaign for support of this work. We thank Dr. Danielle Gray for X-ray analysis of compound AA8. R.J.R. dedicates to Professor Robert M. Williams on the occasion of his 60th birthday.



Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201308743.

^[*] Dr. R. J. Rafferty, R. W. Hicklin, K. A. Maloof, Prof. P. J. Hergenrother Department of Chemistry University of Illinois at Urbana-Champaign 261 RAL, Box 36-5, 600 S. Mathews, Urbana, IL 61801 (USA) E-mail: hergenro@illinois.edu

Scheme 1. Overview of the construction of diverse scaffolds from abietic acid (AA).

caulkings.^[38] Because of its abundance and commercial utility, substantial information exists on the chemical reactivity of AA, resulting from structural elucidation studies,^[39,40] and its use as the starting point in the construction of several natural products.^[41–45] AA is well-suited for diversification through CtD given its fused polycyclic ring system and olefinic functional groups, allowing the selective manipulation of individual rings and multiple ring distortions, enabling the facile creation of highly diverse scaffolds from AA (Scheme 1, AA1–AA11).

The C- and B-ring olefins of AA are synthetic handles that provide entry points to ring distortion. As shown in Scheme 1 for the construction of AA1–AA11 and described below, the differential and systematic manipulation of these olefins allows the creation of multiple scaffolds through the expansion, contraction, fusion, cleavage, and aromatization of rings.

Of the 84 compounds described in this manuscript (all of which are shown in Supporting Figure 1), seven previously reported compounds (specifically AA1,^[46] AA4 and AA6,^[47] AA12,^[46] AA13,^[43] and AA17–18^[43,48,49]) were synthesized, the other 77 compounds reported herein are new compounds.

Dihydroxylation of AA is known to occur with high regioand diastereoselectivity to afford AA12, [46] effectively differentiating the B- and C-ring olefins (Scheme 2). Methylation of free acid AA12 generates methyl ester AA13, which readily undergoes oxidative ring cleavage with lead tetraacetate in benzene to afford AA1.

AA1, which contains three carbonyl functionalities, is a key starting point for creation of structurally diverse scaffolds through ring fusion and intramolecular aldol condensation (Scheme 3). Wittig olefination of AA1 with excess Ph₃PCH₃Br gives rise to triene AA14 (Scheme 3). The conjugated diene of AA14 is well suited for Diels-Alder ring fusion and readily affords AA2 with exclusive facial selectivity. An intramolecular aldol condensation of AA1 was utilized to achieve a formal ring contraction of the AA scaffold. This was accomplished through exposure of AA1 to potassium tert-butoxide in THF to furnish diene AA3. Subjecting AA1 to lithium hydroxide produces the free acid, which in turn can be amidated under PyBOP coupling conditions with benzyl amine to afford AA15. Reductive amination of AA15 with p-methoxybenzylamine and sodium cvanoborohydride furnishes amine AA16.

Scheme 2. Oxidative cleavage of the C ring to product AA1. pyr = pyridine.

Scheme 3. Conversion of AA1 to bi-, tri-, and tetracyclic scaffolds. PyBOP = 1-benzotriazolyloxy-tris(pyrollidino) phosphonium.

Further modifications of the C ring were accomplished by aromatization and oxidation. The methyl ester of abietic acid (AA17) was aromatized using reported procedures to furnish AA18 as a mixture of isomers, which was elaborated to AA4 through a Swern oxidation (Scheme 4a). [46] The C-ring olefin could also be readily converted to a cyclic carbonate. Treatment of diol AA13 with triphosgene gave access to cyclic



Scheme 4. a) Aromatization of the C ring, and b) ring fusion onto the diol of the C ring.

carbonate AA5 (Scheme 4b). Both AA4 and AA5 contain synthetic handles for manipulation of the B ring, as described below.

Ring expansion of AA4 through a Baeyer–Villiger oxidation gives AA6 (Scheme 5). Ring cleavage of lactone AA6 with lithium aluminum hydride provides triol AA7.

Scaffold AA5 can be used to construct several types of products from distortion of the B ring through ring fusion, ring cleavage, and ring expansion (Scheme 6). Cyclopropanation of AA5 using aqueous sodium hydroxide and tetrabutylammonium bromide in bromoform provides dibromocyclopropane AA8, whose stereochemistry was confirmed by X-ray crystallography (Scheme 6). Ring expansion of the

Scheme 5. Ring expansion and cleavage of AA4.

B ring within AA8 was achieved using silver tetrafluoroborate to afford cycloheptadiene AA9. The vinyl bromide of AA9 is readily diversified through transition metal-catalyzed cross coupling reactions (see below).

Treatment of AA5 with osmium tetraoxide gives rise to AA19, which upon treatment with lead tetraacetate produces ring cleavage product AA20 (Scheme 6). The aldehyde in AA20 was utilized in the construction of two compounds (AA10 and AA11). Subjecting AA20 to potassium *tert*-butoxide results in formal ring contraction through intramolecular aldol addition, providing AA10. Alternatively, selective reduction of AA20 with sodium borohydride affords primary alcohol AA11.

As AA1–AA20 visually appear to be diverse, we have applied a similarity matrix to quantify the structural diversity of the collection. For this, Tanimoto similarity coefficients (extended-connectivity fingerprints: ECFP_6) were calculated for pairwise combinations of compounds (Supporting Figure 2).^[50,51] In this method,

each atom is analyzed for connectivity patterns within three bonds relative to the starting atom, thus generating a compound fingerprint. Each of these fingerprints is then compared, generating Tanimoto coefficients of similarity. Compounds with high similarity possess Tanimoto coefficients with values closer to one, whereas values closer to zero indicate lower structural similarity. For example, diol AA12 and its methyl ester AA13 differ only by a single methyl group and have a high Tanimoto similarity (0.75, Supporting Figure 2). In contrast, AA7 and AA8 possess very different scaffolds and have a Tanimoto coefficient of 0.09 (Supporting Figure 2). Using this Tanimoto similarity coefficient, library diversity is indicated by low Tanimoto similarity coefficients for most pairwise combinations. Tanimoto similarity analysis suggests considerable structural diversity within AA1-AA20 in comparison to each other and the parent natural product AA (Supporting Figure 2). Analysis of the scaffolds in Scheme 1 (AA1-AA11) provides an average Tanimoto similarity of 0.26, ranging from 0.09 to 0.56. Further analysis of the entire AA set (AA-AA20) shows a similarly low Tanimoto average of 0.27 (range 0.09 to 0.75).

The molecular complexity of the compounds created from AA was quantified using the fraction of sp³-hybridized carbon atoms (Fsp3) and the number of stereogenic centers, two structural features commonly used as surrogates for molecular complexity. [6,52,53] Fsp3 is defined as the number of sp3hybridized carbon atoms divided by the sum of sp³- and sp²hybridized carbon atoms in a given compound. [54] Molecules with low Fsp3 tend to be flat, achiral, and dominated by aromatic rings. Compounds with high Fsp3 have a greater propensity to possess out-of-plane substituents, possibly allowing increased receptor-ligand interactions. [6] Early stage discovery compounds already have a lower average Fsp3 than FDA-approved small-molecule drugs (0.36 vs. 0.47)^[6] and recent reports suggest a non-ideal situation where the Fsp3 of compounds created through medicinal chemistry is declining, [4] increasing the gap in complexity between approved drugs and discovery compounds.

Scheme 6. Reaction of the B-ring olefin of AA5.

The compounds created through CtD (AA1-AA20) have high Fsp3s (average 0.74, range 0.49 to 0.88; Figure 1 A). These values are markedly different from standard screening collections, as represented by the Chembridge MicroFormat screening library (http://chembridge.com) of over 150,000 compounds, the compounds of which possess an average Fsp3 of 0.23.[36,55] For comparison, the average Fsp3 of all FDAapproved small-molecule drugs is 0.47 (Figure 1 A).

The number of stereogenic centers is a complementary metric for molecular complexity, and compounds with multiple stereogenic centers may interact more favorably with

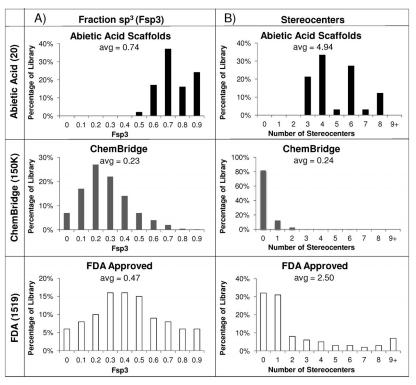


Figure 1. Assessment of structural complexity measured by A) Fsp3 for AA scaffolds (AA1-AA20), ChemBridge MicroFormat collection (150000 compounds), and FDAapproved drugs (1519 compounds, obtained via the Institute for Molecular and Cellular Pharmacology, Nice, France, http://chemoinfo.ipmc.cnrs.fr), and B) number of stereocenters for AA scaffolds (AA1-AA20), ChemBridge MicroFormat collection, and FDA-approved drugs. Discovery Studio Client 2.5 was used for these calculations.

chiral biological receptors. As shown in Figure 1B, the compounds in the AA set all possess multiple stereogenic centers (average 4.94, range 3 to 8). The Chembridge MicroFormat collection has an average of 0.24 stereogenic centers (range 0 to 6). FDA-approved drugs (derived from both synthetic and natural products) have an average of 2.5 stereogenic centers with a range of 0 to 26 (Figure 1B).

FDA-approved drugs have a wide range of Fsp3 and number of stereogenic centers (Figure 1). Compounds from traditional screening libraries (e.g., Chembridge MicroFormat) nicely cover the area of the histograms populated by compounds with low Fsp3 (0-0.5) and few (0-2) stereogenic centers. In a contrasting and complementary fashion, compounds created through CtD have very different Fsp3 and stereogenic properties, with Fsp3 of around 0.5-1, and number of stereogenic centers of approximately 3-8 (Figure 1).

Numerous compounds poised for rapid functional group diversification were created during the course of our synthetic efforts (Scheme 1). Multiple derivatives can be constructed from such compounds, exemplified by the assembly of five single-point-derivative libraries (Supporting Scheme 1).

Single-point derivatives of AA9 were accessed through Suzuki cross-coupling reactions utilizing the vinyl bromide functionality. Coupling with six commercially available bor-

> onic acid pinacol esters resulted in derivatives AA21-AA26 (Supporting Scheme 1 A). Diversification of the peripheral aldehyde of AA10 was accomplished through reductive amination in the presence of various amines to provide AA27-AA30 in modest yields (Supporting Scheme 1B).

> Multiple single-point-derivative libraries can be accessed along a synthetic route by diversification of a suitable intermediate that is then carried through the remaining sequence. For example, derivatization of AA12 through esterification or amidation provided seven single-point derivatives (AA31a-g, Supporting Scheme 1 C). Subjecting these derivatives to oxidative cleavage afforded a second single-point-derivative library (AA32a-g). Exposure of AA32a-g to potassium tert-butoxide resulted in a library consisting of AA33a-g, obtained through intramolecular aldol condensations. Through this approach, 21 compounds (20 new compounds plus AA15) with three distinct scaffolds were expediently accessed.

> To further illustrate the robustness of this methodology in the generation of large numbers of complex compounds, a threepoint derivative library based upon scaffold AA2 was constructed (Supporting Scheme 2). Amidation/esterification, Wittig olefination, and Diels-Alder reactions were chosen for the construction of this library because of their speed, efficiency, and broad substrate scope. Wittig olefinations were performed on

AA1, AA15, and AA32g with three different triphenylphosphine reagents to furnish nine dienes (AA14, AA34a,b, AA35a-c, and AA36a-c). Ring fusion through Diels-Alder cycloaddition was employed as the final step in this sequence to efficiently provide AA2 and 26 variants (AA37-AA45).

Many natural products contain complex scaffolds with rigid ring systems, multiple stereogenic centers, and functional groups suitable for scaffold diversification through ring distortion. Herein, the fused tricyclic core of AA was readily transformed into 20 structurally diverse scaffolds through manipulation of the B- and C-ring olefins. In the course of these studies, several powerful ring distortion reactions were employed including the Diels-Alder cycloaddition (AA2),

227



electrocyclic dihalocyclopropane opening/elimination (AA9), and aldol addition (AA10).

The scaffolds produced through the CtD method are structurally diverse from one another and AA as assessed by Tanimoto similarity (average 0.26). The complexity of these compounds is significantly greater (average Fsp3 0.74, average number of stereogenic centers 4.94) than a typical commercial screening collection (Chembridge MicroFormat, average Fps3 0.23, average number of stereocenters 0.24); thus, compounds created through CtD are complementary to traditional synthetic screening libraries. Furthermore, these scaffolds are well suited for functional-group diversification, as illustrated with the construction of five single-point-derivative libraries and one three-point derivative library. Importantly, AA1–AA20 have been synthesized on 25–3000 mg scale, and library compounds (AA21–AA45) on 3–50 mg scale.

Natural products are valuable starting points for the creation of a multitude of complex and diverse chemical entities. The abundance of AA and the efficiency of ring distortion allowed the preparation of every compound in the 84-membered AA set in multimilligram quantities; the structures of all 84 compounds are shown in Supporting Figure 1. We anticipate that the compounds disclosed herein, and others created by analogous efforts on different natural products, will find utility in a wide variety of biological screens.

Received: October 7, 2013

Published online: November 24, 2013

Keywords: abietic acid \cdot high-throughput screening \cdot molecular diversity \cdot natural products \cdot ring distortion

- D. C. Swinney, J. Anthony, Nat. Rev. Drug Discovery 2011, 10, 507.
- [2] J. T. Heeres, P. J. Hergenrother, Chem. Soc. Rev. 2011, 40, 4398.
- [3] E. J. Perola, J. Med. Chem. 2010, 53, 2986.
- [4] W. P. Walters, J. Green, J. R. Weiss, M. A. Murcko, J. Med. Chem. 2011, 54, 6405.
- [5] P. D. Leeson, S. A. St-Gallay, *Nat. Rev. Drug Discovery* **2011**, *10*, 749
- [6] F. Lovering, J. Bikker, C. Humblet, J. Med. Chem. 2009, 52, 6752.
- [7] K. T. Flaherty, U. Yasothan, P. Kirkpatrick, Nat. Rev. Drug Discovery 2011, 10, 811.
- [8] S. Dandapani, L. Marcaurelle, Nat. Chem. Biol. 2010, 6, 861.
- [9] D. V. Ferraris, J. Med. Chem. 2010, 53, 4561.
- [10] A. Domling, Curr. Opin. Chem. Biol. 2008, 12, 281.
- [11] P. D. Grivas, H. Kiaris, A. G. Papavassiliou, *Trends Mol. Med.* 2011, 17, 537.
- [12] R. O'Shea, H. E. Moser, J. Med. Chem. 2008, 51, 2871.
- [13] L. L. Silver, Clin. Microbiol. Rev. 2011, 24, 71.
- [14] D. J. Payne, M. N. Gwynn, D. J. Holmes, D. L. Pompliano, Nat. Rev. Drug Discovery 2007, 6, 29.
- [15] D. J. Newman, G. M. Cragg, J. Nat. Prod. 2012, 75, 311.
- [16] J. Liu, J. D. Farmer, W. S. Lane, J. Friedman, I. Weissman, S. L. Schreiber, *Cell* 1991, 66, 807.
- [17] M. A. Jordan, L. Wilson, Nat. Rev. Cancer 2004, 4, 253.
- [18] M. D'Incalci, C. M. Galmarini, Mol. Cancer Ther. 2010, 9, 2157.
- [19] R. H. Baltz, Curr. Opin. Pharmacol. 2008, 8, 557.
- [20] L. L. Silver, Clin. Microbiol. Rev. 2011, 24, 71.

- [21] W. R. Galloway, A. Isidro-Llobet, D. R. Spring, *Nat. Commun.* 2010, 1, 80.
- [22] S. L. Schreiber, Science 2000, 287, 1964.
- [23] M. R. Spaller, M. T. Burger, M. Fardis, P. A. Bartlett, *Curr. Opin. Chem. Biol.* **1997**, *1*, 47.
- [24] J. Cui, J. Hao, O. A. Ulanovskaya, J. Dundas, J. Liang, S. A. Kozmin, Proc. Natl. Acad. Sci. USA 2011, 108, 6763.
- [25] O. C. CJ, H. S. Beckmann, D. R. Spring, Chem. Soc. Rev. 2012, 41, 4444.
- [26] K. M. O'Connell, H. S. Beckmann, L. Laraia, H. T. Horsley, A. Bender, A. R. Venkitaraman, D. R. Spring, *Org. Biomol. Chem.* 2012, 10, 7545.
- [27] E. Comer, H. Liu, A. Joliton, A. Clabaut, C. Johnson, L. B. Akella, L. A. Marcaurelle, *Proc. Natl. Acad. Sci. USA* 2011, 108, 6751
- [28] S. Wetzel, R. S. Bon, K. Kumar, H. Waldmann, Angew. Chem. 2011, 123, 10990; Angew. Chem. Int. Ed. 2011, 50, 10800.
- [29] B. C. Goess, R. N. Hannoush, L. K. Chan, T. Kirchhausen, M. D. Shair, J. Am. Chem. Soc. 2006, 128, 5391.
- [30] H. E. Pelish, N. J. Westwood, Y. Feng, T. Kirchhausen, M. D. Shair, J. Am. Chem. Soc. 2001, 123, 6740.
- [31] B. R. Balthaser, M. C. Maloney, A. B. Beeler, J. A. Porco, Jr., J. K. Snyder, *Nat. Chem.* **2011**, *3*, 969.
- [32] K. J. Frankowski, V. Setola, J. M. Evans, B. Neuenswander, B. L. Roth, J. Aube, Proc. Natl. Acad. Sci. USA 2011, 108, 6727.
- [33] C. Aquino, M. Sarkar, M. J. Chalmers, K. Mendes, T. Kodadek, G. C. Micalizio, *Nat. Chem.* 2012, 4, 99.
- [34] J. Aube, Nat. Chem. 2012, 4, 71.
- [35] H. Lachance, S. Wetzel, K. Kumar, H. Waldmann, J. Med. Chem. 2012, 55, 5989.
- [36] R. W. Huigens III, K. C. Morrison, R. W. Hicklin, T. A. Flood, Jr., M. F. Richter, P. J. Hergenrother, *Nat. Chem.* 2013, 5, 195.
- [37] W. H. Sauer, M. K. Schwarz, J. Chem. Inf. Comput. Sci. 2003, 43, 987.
- [38] H. H. Zeiss, Chem. Rev. 1948, 42, 163.
- [39] L. Ruzicka, R. G. Bacon, Helv. Chim. Acta 1936, 19, 1542.
- [40] H. Wienhaus, W. Sandermann, Chem. Ber. Dtsch. Chem. Ges. 1936, 69, 2202.
- [41] H. Okawara, H. Nakai, M. Ohno, Tetrahedron Lett. 1982, 23, 1087.
- [42] H. Koyama, Y. Kaku, M. Ohno, Tetrahedron Lett. 1987, 28, 2863.
- [43] E. Alvarez-Manzaneda, R. Chahboun, F. Bentaleb, E. Alvarez, M. A. Escobar, S. Sad-Diki, M. J. Cano, I. Messouri, *Tetrahedron* 2007, 63, 11204.
- [44] B. Zhou, X. M. Li, H. J. Feng, Y. C. Li, Tetrahedron 2010, 66, 5396.
- [45] J. S. Yadav, G. Baishya, U. Dash, Tetrahedron 2007, 63, 9896.
- [46] E. Alvarez-Manzaneda, R. Chahboun, E. Cabrera, E. Alvarez, R. Alvarez-Manzaneda, M. Lachkar, I. Messouri, *Tetrahedron Lett.* 2007, 48, 989.
- [47] E. Alvarez-Manzaneda, R. Chahboun, E. Alvarez, R. Alvarez-Manzaneda, P. E. Munoz, F. Jimenez, H. Bouanou, *Tetrahedron* 2011, 67, 8910.
- [48] K. Krohn, E. Budianto, U. Florke, B. M. Hausen, *Liebigs Ann. Chem.* 1992, 911.
- [49] R. C. Cambie, L. H. Mitchell, P. S. Rutledge, Aust. J. Chem. 1998, 51, 1167.
- [50] D. Rogers, M. Hahn, J. Chem. Inf. Model. 2010, 50, 742.
- [51] D. J. Rogers, T. T. Tanimoto, Science 1960, 132, 1115.
- [52] P. Selzer, H. J. Roth, P. Ertl, A. Schuffenhauer, Curr. Opin. Chem. Biol. 2005, 9, 310.
- [53] F. Lovering, Med. Chem. Commun. 2013, 4, 515.
- [54] A. Yan, J. Gasteiger, J. Chem. Inf. Comput. Sci. 2003, 43, 429.
- [55] M. Feher, J. M. Schmidt, J. Chem. Inf. Comput. Sci. 2003, 43, 218.