Combinatorial Chemistry Hot Paper

DOI: 10.1002/anie.201307786

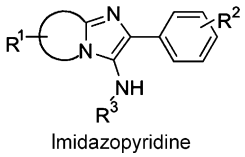
Combining On-Chip Synthesis of a Focused Combinatorial Library with Computational Target Prediction Reveals Imidazopyridine GPCR Ligands**

*Michael Reutlinger, Tiago Rodrigues, Petra Schneider, and Gisbert Schneider**

Abstract: *Using the example of the Ugi three-component reaction we report a fast and efficient microfluidic-assisted entry into the imidazopyridine scaffold, where building block prioritization was coupled to a new computational method for predicting ligand–target associations. We identified an innovative GPCR-modulating combinatorial chemotype featuring ligand-efficient adenosine $A_{1/2B}$ and adrenergic $\alpha_{1A/B}$ receptor antagonists. Our results suggest the tight integration of microfluidics-assisted synthesis with computer-based target prediction as a viable approach to rapidly generate bioactivity-focused combinatorial compound libraries with high success rates.*

The fast pace of drug-discovery programs is supported by high-throughput screening campaigns to identify new chemical entities, where the underlying compound collections used for screening benefit from combinatorial libraries with lead- and drug-like properties.[1] While numerous synthesis protocols are available, a reliable assessment of potential macromolecular targets of these compounds is desirable for the compilation of bioactivity-focused combinatorial libraries. For the Ugi four- and three-component reactions,[2] which have shown robustness in producing both model compounds and drug candidates,[3,4] we report a fast and efficient microfluidic-assisted entry into the imidazopyridine scaffold, coupled to a new computational prediction method for "deorphaning" ligand–target associations. We identified an innovative GPCR-modulating combinatorial chemotype featuring ligand-efficient adenosine $A_{1/2B}$ and adrenergic $\alpha_{1A/B}$ receptor antagonists (GPCR = G-protein coupled receptor). Our results suggest that the tight integration of microfluidics-assisted synthesis with computer-based target prediction is a viable approach to rapidly generate bioactivity-focused combinatorial compound libraries with high success rates.

Imidazopyridines may be considered to be a privileged scaffold given their diverse range of macromolecular drug targets.[5–8] While entry into this chemotype through an Ugi three-component reaction has been reported,[5–7] these methods do not allow for the quick assembly of combinatorial libraries and scaling up. Therefore, our initial efforts focused on developing a robust and scalable process in flow using a continuous synthesis system equipped with low-pressure, pulsation-free syringe pumps. The setup included a 3-2-way solenoid valve to allow for automated cycles of building block filling and dispensing. The amine and benzaldehyde components were dissolved in ethanol, together with perchloric acid, while the isocyanide component was pumped independently. The concentrations of the stock solutions were adjusted to afford the desired final concentrations in the microreactor. A borosilicate DeanFlow chip with a total volume of 5 μL and a zig-zag mixing zone was used as the primary reactor (Figure 1 A). Alternatively, we used a KombiMix chip with a reaction volume of 13 μL (Figure 1 B). The protocol was then scripted with Cetoni Q_{mix} Elements software to automate all steps, including the washing of the microfluidic channels.



Imidazopyridine

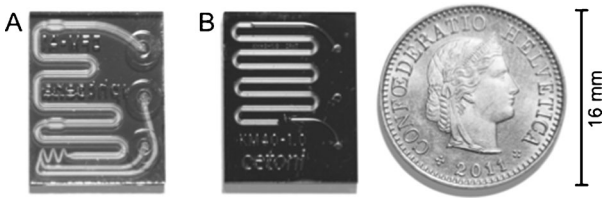


Figure 1. DeanFlow (A) and KombiMix (B) microreactor chips.

In an initial screening of the reaction conditions we performed sequential and automated syntheses of compound **1**. Conversion rates were derived from ^{1}H NMR spectra (Figure 2 A,B). In the first instance we investigated optimal flow rates and reaction temperatures, using 10 mol% of catalyst and a final concentration of each building block equal to 0.3 M, as described previously.[6] Generally, the reactions gave better results at lower temperatures (30 and 70 °C) than at 170 and 200 °C. Additionally, we conducted control reactions in glassware at room temperature and 30 °C for two hours, and measured conversions of 73% and 80%, respectively. The results pinpoint the usefulness of a microreactor, both for improving conversions and drastically shortening reaction times. Reactions carried out under higher flow rates (30 and 60 μL s^{-1}) gave poorer results than

[*] M. Reutlinger, Dr. T. Rodrigues, Dr. P. Schneider, Prof. Dr. G. Schneider
Swiss Federal Institute of Technology (ETH)
Department of Chemistry and Applied Biosciences
Wolfgang-Pauli-Strasse 10, 8093 Zurich (Switzerland)
E-mail: gisbert.schneider@pharma.ethz.ch
Dr. P. Schneider, Prof. Dr. G. Schneider
inSili.com GmbH
Segantinisteig 3, 8049 Zurich (Switzerland)

[**] This research was financially supported by a research grant from the OPO Foundation, Zurich. GPCR = G-protein coupled receptor.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201307786.

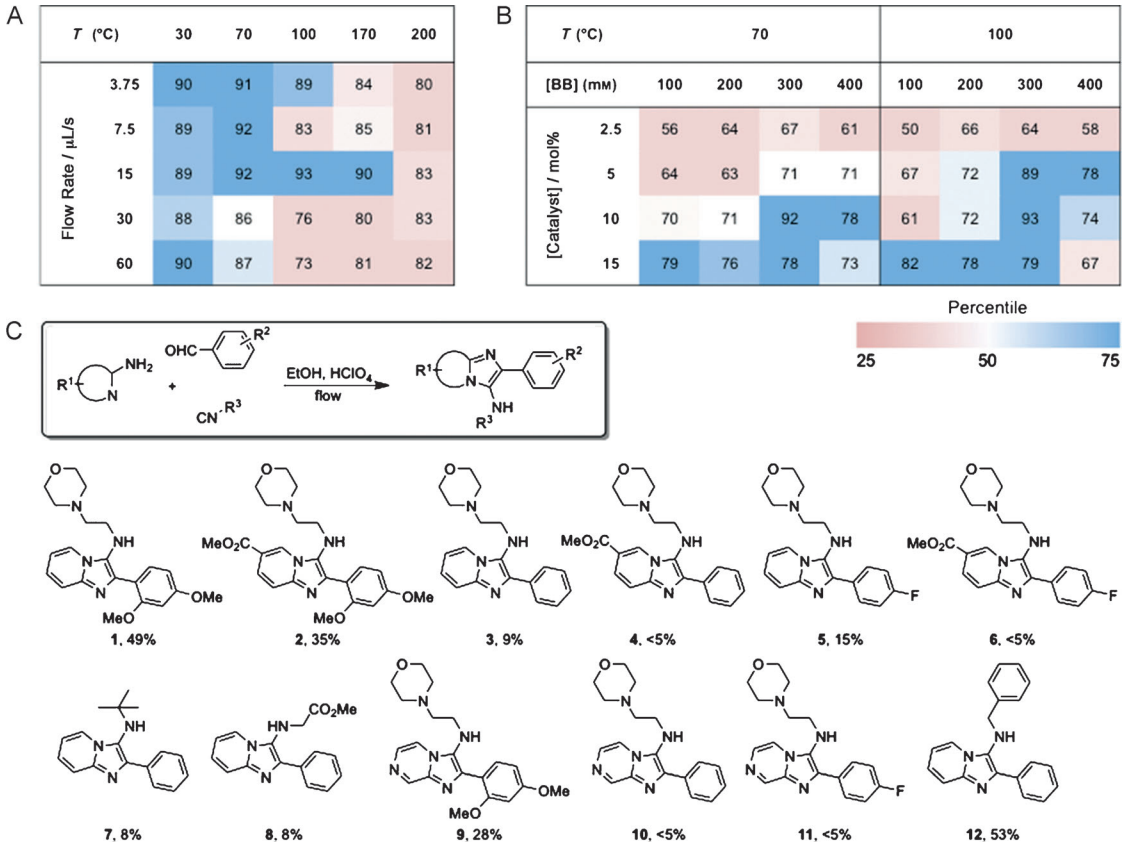


Figure 2. Synthesis of imidazopyridines in flow: A) Screening of optimal flow rate and temperature (*T*), at constant catalyst loading (10 mol%) and building block (BB) concentration (0.3 M); B) Screening of optimal catalyst loading and building block concentration, at fixed flow rate (15 $\mu L s^{-1}$) and temperature (70 and 100 °C); C) Focused library synthesized in the present study and yields of isolated products.

their 3.75 $\mu L s^{-1}$ and 7.5 $\mu L s^{-1}$ counterparts, possibly due to shorter residence times in the reactor chip. We observed the highest conversion at intermediate temperatures (70 and 100 °C). Interestingly, at 70 °C the reaction appears to tolerate a wide range of flow rates, while at 100 °C a rate of 15 $\mu L s^{-1}$ is preferable.

Having determined the best binary combination of temperature and flow rate, we screened for the ideal catalyst loading and final concentration of building blocks and found these to be 10 mol% and 0.3 M, respectively (Figure 2B). Comparable conversion rates were obtained in a microwave procedure[6] and the setup described herein (94% vs. 93%, respectively). Of note, these results were obtained using a lower reaction temperature in the flow system (100 °C in flow vs. 170 °C in the microwave reactor) and shorter reaction times (0.3 s in flow vs. 15 min in the microwave reactor). Finally, the optimized reaction conditions were compared in the DeanFlow and KombiMix microreactors. While the conversion of **1** in the DeanFlow chip was 93%, an 88% conversion was observed in the KombiMix chip.

With these results in hand we synthesized a small focused library of imidazopyridines **1**–**12** (Figure 2C) using the DeanFlow reactor chip, and predicted potential biological targets with Gaussian process regression models, which were constructed from 469 drug targets that are annotated in the ChEMBL database (version 14).[9] Given a query compound, the computer model predicts p*Affinity* values for each target, which goes beyond related computational tools.[10] Furthermore, to ensure meaningful, nontrivial, and high-value predictions we calculated the Mahalanobis distance (MD) of the predicted values to the predictions made for a large collection of randomly selected molecules. Here, we considered only drug targets for which we obtained p*Affinity* > 5.5 and MD > 0.5 standard deviations. With these mildly restrictive criteria we predicted an average of 18 targets per compound. Basically due to the low p*Affinity* bound, this number exceeds other theoretical considerations and experimental findings reporting between two and ten targets per drug, depending on the target class.[11] We obtained an average of four targets per imidazopyridine compound with the more conservative boundaries p*Affinity* > 6 and MD > 1. Keeping the permissive estimate we selected a total of 41 targets with high p*Affinity* predictions for further study. For these targets the model yielded favorable cross-validated accuracies of $Q^2 = 0.68 \pm 0.10$, $MAE = 0.65 \pm 0.11$ and $BEDROC = 0.67 \pm 0.15$ (all values mean ± standard deviation).[12]

We finally selected five targets based on majority predictions for the whole library, potential pharmaceutical interest, and assay availability. p*Affinity* values were in the micromolar range (Table 1); even though the predicted variance was high. This observation emphasizes the potential

Table 1: Summary of results for selected compounds **1**, **3**, **5**, **7**, **8**, **9**, and **12**.

	Target	Predicted p*Affinity*	Mahalanobis distance	Experimental pK_i or % binding	*LE*[a]	*LLE*[b]	*SILE*[c]
1	α_{1A}[d]/PDE10A[e]	5.7/5.7	0.7/0.8	$<4/<4$	–	–	–
3	α_{1B}[f]	6.2	2.4	5.6	0.33	3.46	3.04
5	α_{1A}/A_{2B}[g]	5.8/6.5	0.7/2.4	5.4/5.2	0.30/0.29	3.07/2.86	2.87/2.76
7	α_{1B}	6.1	2.0	5.7	0.40	1.74	3.23
8	A_1[h]	5.7	3.2	$>80\%$[i]	–	–	–
9	A_{2B}/PDE10A	6.4/5.8	2.6/1.7	$<4/<4$	–	–	–
12	A_1	6.0	3.3	$>80\%$[i]	–	–	–

[a] Ligand efficiency. [b] Lipophilic ligand efficiency. [c] Size-independent ligand efficiency. [d] Adrenergic α_{1A} receptor. [e] Phosphodiesterase 10A. [f] Adrenergic α_{1B} receptor. [g] Adenosine A_{2B} receptor. [h] Adenosine A_1 receptor. [i] Radioligand assay; activity values are averaged from two measurements.

novelty of the scaffold compared to known ligands in the ChEMBL database. In fact, to the best of our knowledge, imidazopyridines with this framework have not been reported as adenosine or adrenergic receptor ligands.[20]

Having predicted potential macromolecular targets for all synthesized compounds, we tested those compounds for which we had obtained robust p*Affinity* predictions. For one of the prominent targets, phosphoinositide 3-kinase, activity had previously been reported for the underlying imidazopyridine scaffold,[6] which corroborated the prediction. As a proof-of-concept, we then explored a range of predicted GPCR targets aiming at the discovery of a new activity island in chemical space. In radioligand displacement assays probing the direct ligand–receptor binding and in cell-based functional activity assays, 71% of the compounds were found to be active as predicted (Table 1). More specifically, compounds **3** and **7** presented antagonistic K_i values of 2–3 μM, respectively, against the adrenergic α_{1B} receptor, while compound **5** showed similar low micromolar antagonistic potency against the adrenergic α_{1A} and adenosine A_{2B} receptors. Compounds **8** and **12** turned out to be potent direct ligands of the A_1 receptor (84% and 89% binding at 100 μM, respectively), but were inactive in the functional cell-based assay. Additional tests will be required to determine selectivity profiles in a full GPCR panel screen.

Several quality indices have been suggested to guide hit prioritization in drug discovery.[13] Accordingly, our compounds fully qualify as lead structure candidates (Table 1). For example, compound **7** is a scarcely decorated, yet highly ligand-efficient chemical entity ($LE = 0.40$; $SILE = 3.23$) which might justify development as an adrenergic α_{1A} receptor antagonist. On the other hand, although less efficient than **7**, compound **3** presents a better balance between affinity and computed log*P*(o/w) ($LLE = 3.46$ vs. 1.74). Most importantly, the leads presented herein are dissimilar to their nearest neighbors from the training data (structural similarity Tanimoto = 0.16–0.30, Table S1) and would likely not have been selected using straightforward substructure-based similarity searching.

Altogether, our chemistry-driven approach to the design of a target-focused combinatorial library, in an expeditious and efficient manner, led to the identification of a molecular framework targeting four GPCRs. The results highlight the imidazopyridine scaffold as a privileged motif and demonstrate how the integration of emerging technologies in drug discovery, such as on-chip synthesis and computational target prediction, may advance hit and lead identification in chemical biology and molecular medicine. In light of recent advances in lab-on-a-chip technologies,[14] one could even envisage a fully automated hit-finding automaton that integrates computational target prediction and building block selection for the microfluidic-assisted synthesis and testing of candidate compounds.

Experimental Section

Computations. For training the Gaussian process models[15] we used the ChEMBL database (version 14) containing 1213242 distinct compounds with 10129256 bioactivities for 9003 targets.[9] Protein targets with fewer than 200 annotated human bioactivities were excluded. All activity end-points were standardized to p*Affinity* = $-\log_{10}$(*activity*). The final affinity data set consisted of 209293 compounds with 431313 bioactivities for 469 human targets. Post-processing was conducted using Python (http://www.python.org) and Knime v.2.6.0.[16] Molecular structures were standardized using the "wash" function in MOE 2012.10 (The Chemical Computing Group Inc., Montreal, Canada); log*P*(o/w) was calculated with MOE. Two different molecular descriptors were calculated for each compound: topological pharmacophores (CATS2, 0–9 bonds, type-sensitive scaling),[6] and an ECFP-like topological circular fingerprint (Morgan fingerprint, *radius* = 4, 2048 bit; RDKit: http://www.rdkit.org).[17] Predictive models were implemented using Matlab R2012b (The MathWorks Inc., Natick, USA) and the GPML toolbox v3.1 (http://www.gaussianprocess.org). We assessed prediction quality by tenfold stratified cross-validation (cross-validated squared correlation coefficient, Q^2; mean absolute error, MAE). The Boltzmann-enhanced discrimination of ROC (BEDROC; $\alpha = 56$, top 3% contribute 80% to the score) was used to quantify the early enrichment performance.[18] We used the lower confidence-bound p*Affinity* estimate throughout this study: prediction $= \mu* - \sigma^2*$, where μ_* is the model's predictive mean and σ^2_* the predictive variance. To distinguish from random predictions we calculated the Mahalanobis distance of an activity prediction: MD(*prediction*) = (*prediction*$-\mu_r$)/σ_r, where μ_r and σ_r are the mean and standard deviation of a randomized predictive distribution. The background consisted of 50000 randomly selected molecules from ChemDB.[19]

Synthesis. Stock solutions of building blocks were prepared in ethanol. The amine and aldehyde components were premixed, and perchloric acid was added. Two independent syringe pumps delivered the amine/benzaldehyde/perchloric acid solution and the isocyanide solution at suitable flow rates. The reaction chamber containing the microchip was heated at different temperatures and the crude product was collected in a vial. The crude mixtures were purified by preparative HPLC (acetonitrile/H_2O + 0.1% formic acid in each solvent) using a gradient of 30–95% or 5–50% acetonitrile over 16 min. Microfluidics hardware and the Q_{mix} Elements software were obtained from Cetoni (Korbussen, Germany). Microwave synthesis was performed in a Biotage Initiator (Uppsala, Sweden) in 1–2 mL vials, as described.[6]

Testing. Activity determinations were performed by Cerep (Le Bois l'Evêque, 86600 Celle l'Evescault, France) on a fee-for-service basis. For details see the Supporting Information.

Received: September 4, 2013
Published online: November 26, 2013

Keywords: combinatorial chemistry · computer chemistry · drug design · microfluidics · multicomponent reactions

[1] a) R. Macarron, M. N. Banks, D. Bojanic, D. J. Burns, D. A. Cirovic, T. Garyantes, D. V. S. Green, R. P. Hertzberg, W. P. Janzen, J. W. Paslay, U. Schopfer, G. S. Sittampalam, *Nat. Rev. Drug Discovery* **2011**, *10*, 188 – 195; b) M. C. Bryan, C. D. Hein, H. Gao, X. Xia, H. Eastwood, B. A. Bruenner, S. W. Louie, E. M. Doherty, *ACS Comb. Sci.* **2013**, *15*, 503 – 511.

[2] I. Ugi, *Angew. Chem.* **1962**, *74*, 9 – 22; *Angew. Chem. Int. Ed. Engl.* **1962**, *1*, 8 – 21.

[3] B. Beck, S. Srivastava, K. Khoury, E. Herdtweck, A. Dömling, *Mol. Diversity* **2010**, *14*, 479 – 491.

[4] C. Kalinski, H. Lemoine, J. Schmidt, C. Burdack, J. Kolb, M. Umkehrer, G. Ross, *Synthesis* **2008**, 4007 – 4011.

[5] H. Bienaymé, K. Bouzid, *Angew. Chem.* **1998**, *110*, 2349 – 2352; *Angew. Chem. Int. Ed.* **1998**, *37*, 2234 – 2237.

[6] M. Reutlinger, C. P. Koch, D. Reker, N. Todoroff, P. Schneider, T. Rodrigues, G. Schneider, *Mol. Inf.* **2013**, *32*, 133 – 138.

[7] M. Hieke, C. B. Rödl, J. M. Wisniewska, E. Buscató, H. Stark, M. Schubert-Zsilavecz, D. Steinhilber, B. Hofmann, E. Proschak, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 1969 – 1975.

[8] a) S. Meister, D. M. Plouffe, K. L. Kuhen, G. M. C. Bonamy, T. Wu, S. W. Barnes, S. E. Bopp, R. Borboa, A. T. Bright, J. Che, S. Cohen, N. V. Dharia, K. Gagaring, M. Gettayacamin, P. Gordon, T. Groessl, N. Kato, M. C. S. Lee, C. W. McNamara, D. A. Fidock, A. Nagle, T.-g. Nam, W. Richmond, J. Roland, M. Rottmann, B. Zhou, P. Froissard, R. J. Glynne, D. Mazier, A. Chatterjee, T. T. Diagana, E. A. Winzeler, *Science* **2011**, *334*, 1372 – 1377; b) T. Wu, A. Nagle, K. Kuhen, K. Gagaring, R. Borboa, C. Francek, Z. Chen, D. Plouffe, A. Goh, S. B. Lakshminarayana, J. Wu, H. Q. Ang, P. Zeng, M. L. Kang, W. Tan, M. Tan, N. Ye, X. Lin, C. Caldwell, J. Ek, S. Skolnik, F. Liu, J. Wang, J. Chang, C. Li, T. Hollenbeck, T. Tuntland, J. Isbell, C. Fischli, R. Brun, M. Rottmann, V. Dartois, T. Keller, T. Diagana, E. Winzeler, R. Glynne, D. C. Tully, A. K. Chatterjee, *J. Med. Chem.* **2011**, *54*, 5116 – 5130; c) A. T. Baviskar, C. Madaan, R. Preet, P. Mohapatra, V. Jain, A. Agarwal, S. K. Guchhait, C. N. Kundu, U. C. Banerjee, P. V. Bharatam, *J. Med. Chem.* **2011**, *54*, 5013 – 5030; d) J. M. Wisniewska, C. B. Rödl, A. S. Kahnt, E. Buscato, S. Ulrich, Y. Tanrikulu, J. Achenbach, F. Rörsch, S. Grösch, G. Schneider, J. Cinatl, Jr., E. Proschak, D. Steinhilber, B. Hofmann, *Biochem. Pharmacol.* **2012**, *83*, 228 – 240; e) S. Nordhoff, H. Deppe, U. Abel, A. Feurer, I. Ott, G. Metz, EP1974729 A1, **2008**; f) B. E. Evans, K. E. Rittle, M. G. Bock, R. M. DiPardo, R. M. Freidinger, W. L. Whitter, G. F. Lundell, D. F. Veber, P. S. Anderson, R. S. L. Chang, V. J. Lotti, D. J. Cerino, T. B. Chen, P. J. Kling, K. A. Kunkel, J. P. Springer, J. Hirshfield, *J. Med. Chem.* **1988**, *31*, 2235 – 2246.

[9] A. Gaulton, L. J. Bellis, A. P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, J. P. Overington, *Nucleic Acids Res.* **2011**, *40*, D1100 – D1107.

[10] a) E. Lounkine, M. J. Keiser, S. Whitebread, D. Mikhailov, J. Hamon, J. L. Jenkins, P. Lavan, E. Weber, A. K. Doak, S. Côté, B. K. Shoichet, L. Urban, *Nature* **2012**, *486*, 361 – 367; b) J. Besnard, G. F. Ruda, V. Setola, K. Abecassis, R. M. Rodriguiz, X. P. Huang, S. Norval, M. F. Sassano, A. I. Shin, L. A. Webster, F. R. Simeons, L. Stojanovski, A. Prat, N. G. Seidah, D. B. Constam, G. R. Bickerton, K. D. Read, W. C. Wetsel, I. H. Gilbert, B. L. Roth, A. L. Hopkins, *Nature* **2012**, *492*, 215 – 220.

[11] a) A. A. Antolín, J. Mestres, *Therapeutic Targets*, Wiley, New York, pp. 309 – 326; b) Y. Hu, J. Bajorath, *AAPS J.* **2013**, *15*, 808 – 815.

[12] a) A. Tropsha, *Mol. Inf.* **2010**, *29*, 476 – 488; b) A. Nicholls, *J. Comput.-Aided Mol. Des.* **2008**, *22*, 239 – 255; c) K. Roy, I. Mitra, *Comb. Chem. High Throughput Screening* **2011**, *14*, 450 – 474.

[13] a) M. M. Hann, G. M. Keserü, *Nat. Rev. Drug Discovery* **2012**, *11*, 355 – 365; b) G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan, A. L. Hopkins, *Nat. Chem.* **2012**, *4*, 90 – 98.

[14] a) P. N. Nge, C. I. Rogers, A. T. Woolley, *Chem. Rev.* **2013**, *113*, 2550 – 2583; b) D. Lombardi, P. S. Dittrich, *Expert Opin. Drug Discovery* **2010**, *5*, 1081 – 1094; c) T. Tsukahara, K. Mawatari, T. Kitamori, *Chem. Soc. Rev.* **2010**, *39*, 1000 – 1013.

[15] a) C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, Cambridge, **2006**; b) M. Rupp, T. Schroeter, R. Steri, H. Zettl, E. Proschak, K. Hansen, O. Rau, O. Schwarz, L. Müller-Kuhrt, M. Schubert-Zsilavecz, K. R. Müller, G. Schneider, *ChemMedChem* **2010**, *5*, 191 – 194.

[16] M. R. Berthold, N. Cebron, F. Dill, T. Gabriel, T. Koetter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel, *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL)*, Springer, Berlin, **2007**, pp. 319 – 326.

[17] D. Rogers, M. Hahn, *J. Chem. Inf. Model.* **2010**, *50*, 742 – 754.

[18] J. F. Truchon, C. I. Bayly, *J. Chem. Inf. Model.* **2007**, *47*, 488 – 508.

[19] J. H. Chen, E. Linstead, S. J. Swamidass, D. Wang, P. Baldi, *Bioinformatics* **2007**, *23*, 2348 – 2351.

[20] Chemical Abstracts Service, SciFinder, https://scifinder.cas.org/, Columbus, OH, U.S.A.

