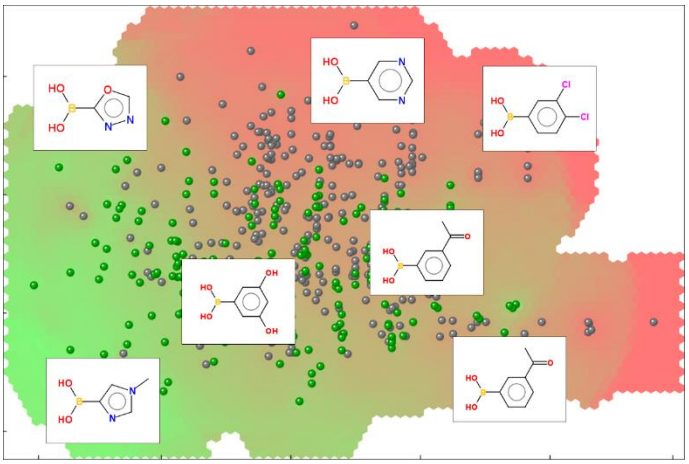
Which boronic acids are used most frequently for synthesis of bioactive molecules?

Peter Ertl, Eva Altmann, Sophie Racine and Odile Decoret
Novartis Institutes for BioMedical Research, Novartis Campus, CH-4056 Basel, Switzerland

Abstract



Boronic acids are essential building blocks used for the synthesis of bioactive molecules, the generation of chemical libraries and the exploration of structure-activity relationships. As a result, more than ten thousand boronic acids are commercially available. Medicinal chemists are therefore facing a challenge; which of them should they select to maximize information obtained by the synthesis of new target molecules. The present article aims to help them to make the right choices. The boronic acids used frequently in the synthesis of bioactive molecules were identified by mining several large molecular and reaction databases and their properties were analyzed. Based on the results a diverse set of boronic acids covering well the bioactive chemical space was selected and is suggested as a basis for library design for the efficient exploration of structure-activity relationships. A Boronic Acid Navigator web tool which helps chemists to make their own selection is also made available at https://bit.ly/boronics.

Keywords:
boronic acids, Suzuki-Miyaura cross coupling, building blocks, Craig plot, cheminformatics, chemical space

1. Introduction

Boronic acids (BAs) are very useful building blocks used in medicinal chemistry for the synthesis of bioactive molecules and the production of chemical libraries enabling rapid exploration of structure-activity relationships (SAR). The major reason why BAs are so popular is the fact that they are used as building blocks in the Suzuki-Miyaura coupling reaction.[1] This reaction was developed for the creation of carbon-carbon bonds between aryl or heteroaryl BAs and aryl or heteroaryl halides or triflates using palladium-based catalysts. Suzuki couplings enable the synthesis of complex molecules with high efficiency and selectivity under mild reaction conditions. The reaction conditions have been extensively studied and are compatible with a broad variety of functional groups, including alcohols, amines, ketones, and esters, among others. The crucial importance of Suzuki cross couplings in medicinal chemistry has been documented by several recent studies. An analysis of reactions reported in medicinal chemistry journals by leading pharmaceutical companies [2] identified Suzuki couplings as the most frequent carbon-carbon forming reaction accounting for 40% of all such reactions. This trend was confirmed in another

study of reactions used in medicinal chemistry by AstraZeneca scientists.[3] An analysis of reactions used in patents also reported a steep increase of Suzuki couplings after 1989, superseding all other carbon-carbon forming reaction types.[4]
A study of parallel libraries performed in AbbVie[5] reported that the Suzuki reaction was the 2nd most used reaction to create libraries (after amide couplings) and BAs used as the most frequent building block after carboxylic acids and amines. BAs are also used in various eco-compatible synthetic techniques, utilizing reaction activation by microwaves, ultrasound, grinding (mechanochemistry) and light. The reactions can be performed in water or other green solvents with phase-transfer catalysis or even in solventless conditions.[6,7] Very recently highly automated capsule-based Suzuki cross couplings have been reported [8] demonstrating the broad interest in this powerful transformation.

Another important reaction using BAs as reagents is Chan-Lam coupling.[9] This is a cross-coupling reaction between an aryl or vinyl halide and an amine or an amide, catalyzed by copper complexes, resulting in the formation of a new carbon-nitrogen bond. Also this reaction can be performed under mild reaction conditions and is compatible with various functional groups.

Several other less known reactions also use BAs as reaction components, for example the Petasis reaction, a multi-component reaction of a vinyl- or aryl- BAs with an amine and a carbonyl derivative to form substituted amines[10], the Liebeskind–Srogl coupling, a general ketone synthesis by the reaction of BAs with thioesters[11] or rhodium catalyzed addition of BAs to alkenes or carbonyl compound.[12] The popularity of all these reactions led to the situation that a large number of BA building blocks is readily available and can be purchased commercially.[13,14]

The importance of BAs is caused, however, not only by their use as synthetic building blocks, but also by application as active agents themselves. BAs can reversibly bind diols, a structural feature that is typical for saccharides, which allows their use as the saccharide sensors, for example in detection of glucose in diabetes patients.[15] The reaction of BAs with diols is used also to generate self-assembled structures in emerging areas of organic material science with applications in gas storage or catalysis.[16] BAs have been also used to facilitate drug and macromolecule delivery, either through incorporation into lipid bilayer for entry via liposomes or through reversible conjugation to a protein.[17] And, last but not least, several BA derivatives are used directly as FDA approved drugs, with others being in development.[18,19] BA derivatives are used as serine protease inhibitors; a prominent example is the proteasome inhibitor Bortezomib for the treatment of multiple myeloma. Other examples include dipeptidyl peptidase-4 inhibitors[20] for the treatment of type 2 diabetes.

In the present study we will focus on the use of BAs as building blocks for synthesis of bioactive molecules. First a medicinal chemistry relevant BAs will be identified by analyzing several data sources commonly used by medicinal chemists. Then these molecules will be characterized by suitable descriptors and the differences between bioactive and average BAs will be explored. And finally, the results will be used to select a subset of BAs that provide maximal structure-activity information in synthesis or library design.

Not only BAs but also their esters are often used as reagents in the Suzuki-Miyaura reaction although they are slightly less reactive than BAs themselves.[21] The most commonly used boronic ester in Suzuki coupling is the pinacol ester, thanks to a combination of its reactivity, stability and ease of preparation.[1] Although in this study we speak mostly about properties and diversity of boronic acids, the conclusions are equally valid for the esters as well.

2. Methodology

2.1 Data sources

BAs from 3 large data sources - commercial compound catalogs, reactions mentioned in patents and building blocks used in synthesis of bioactive molecules - were extracted. All molecules were processed in the same way. Only BAs where the boron atom was connected to the aromatic carbon were processed. All structures were standardized and molecules containing non-standard isotopes and inorganic atoms were discarded. From the rest only BAs with 15 or less atoms were retained. This limit assures good diversity of the final set and at the same time removes too complex and exotic structures. In this study we focused only on the aromatic BAs, since they are clearly the most common building blocks used in the Suzuki reaction. BAs where the $B(OH)_2$ group is connected to a non-aromatic sp2 carbon, although also sometimes used as reagents in the Suzuki coupling, were not covered.

2.1.1. Commercially available boronic acids

The list of BAs available from commercial compound providers was extracted from the PubChem database.[22] PubChem contains data from 919 data sources, 452 of them being chemical vendors.[23] Vendor catalogs that have been updated within the last 3 years and that contain at least 10 thousand molecules (altogether 90 catalogs) were processed and the BAs extracted. This analysis provided 11185 unique BAs, the most common being offered by 61 vendors. 3296 BAs were present in at least 10 catalogs. We did not perform any price analysis at this step, but one can suppose that these common and readily available building blocks are also reasonably priced. This number agrees well with the 4293 BAs with price up to 1000 USD per gram obtained by the analysis of Kalliokoski.[13] The list of commercially available BAs contains, of course, also many rare and quite exotic structures, nearly half of them (5198) being singletons, offered by only 1 vendor. The number of vendors offering particular BAs for sale is also included as a data item in a data table offered for download as Supporting Information.

Number of boronic pinacol esters available commercially is even larger than that of the respective boronic acids (17035 esters were identified by the same procedure as used for BAs). This reflects the greater stability of the esters, particularly those that have a heteroatom in the 2 position or phenyl rings with multiple acceptor substituents.[24]

2.1.2. Boronic acids from patents

BAs that have been used in reactions reported in patents were extracted from United States patent applications and patent grants from years 1976-2013.[25] This data set contains 3.7 million reactions encoded as reaction SMILES. First all reactions were normalized, ions, inorganic components and reactants with valence errors discarded and the atom-atom mapping removed. Only reactions where one of the reactants was a BA were retained. To assure that the reaction indeed describes a Suzuki coupling the following procedure was applied. The chloro or bromo substituted aromatic system in reactants was marked as a potential reaction partner of the BA and the theoretical Suzuki product was generated by merging these 2 reactants. This product was then compared (using canonical SMILES) with the molecule listed in the product side of the reaction SMILES. If both structures were identical this proved that the reaction indeed describes the Suzuki-Miyaura coupling. The BA involved in the reaction was stored and also its reaction

partner was retained to be used later in the validation of reactions used for synthesis of bioactive molecules (see below). From the final collection of these reaction partners the rings that are frequently used as the substrates reacting with BAs were obtained by removing all substituents and also other fused rings. Since as a result of this processing in some cases the generated rings did not have all valences satisfied, they are shown not as complete molecules, but as fragments in Figure 1. The procedure described above identified 18257 unique Suzuki reactions, 1115 BAs and 26 aromatic rings involved as common substrates reacting with the BAs.

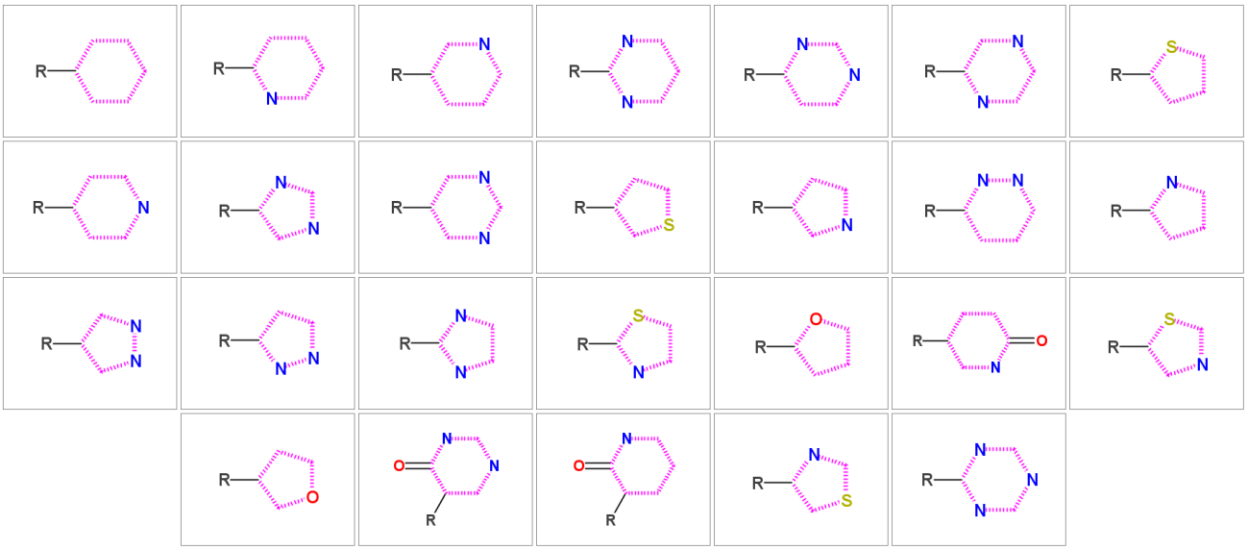


Figure 1. The rings that are used most frequently as substrates reacting with boronic acids in Suzuki-Miyaura coupling. The R represents chlorine or bromine atom, the dashed bonds indicate an aromatic system.

2.1.3. Boronic acids used in the synthesis of bioactive molecules

The BAs that have been used as building blocks in the synthesis of bioactive molecules were obtained by the analysis of molecules from the ChEMBL database.[26] ChEMBL is an indispensable resource for medicinal chemists and cheminformaticians alike, containing in its 32nd release information about 2.3 million molecules, 15 thousand targets and 20 million bioactivity data points extracted from 86 thousand documents, mostly articles in medicinal chemistry journals. Since in the ChEMBL database only the final molecules are stored and not reactions used to prepare them a following procedure had to be used to extract the BA building blocks that have been (probably) used for their synthesis. In the first step molecule series used for exploration of SAR were extracted from medicinal chemistry journals. All molecules in the series had to be reported in a single publication and tested on the same target by the same assay (the same ChEMBL assay ID) and the series had to contain at least 5 molecules. This procedure identified 3453 such series, most of them coming from *Bioorg. Med. Chem. Lett.*, *J. Med. Chem.* and *Eur. J. Med. Chem.* In the next step the series were identified where the derivatives differ in the aromatic substituents connected to another aromatic ring and were therefore possibly synthesized by the Suzuki reaction. All molecules in a series were split on the carbon-carbon bonds connecting these 2 aromatic systems and the R-groups connected to the same core collected. These R-groups represent possibly BA building blocks. Of course, not all connections between 2 aromatic rings

are created by the Suzuki-Miyaura coupling. There are several other reaction types that can be used for this purpose. Particularly molecules containing 5-membered aromatic rings with 2 or 3 heteroatoms (oxazoles, isoxazoles, oxadiazoles, triazoles etc.) may be prepared not by connecting the 2 aromatic rings, but by various heterocyclization reactions[27,28], for example by 1,3-dipolar cycloaddition. To filter out such cases the list of "preferred Suzuki substrates" obtained previously by processing the reactions described in patents (Figure 1) was used. Only the BAs connected to one of these preferred rings were retained. The BAs obtained by this procedure were processed by an additional filtering step and only the structures reported either as reactants in the Suzuki reactions from patents or offered in at least 5 commercial catalogs were retained. Of course, without manually checking all several thousand original literature sources one cannot be sure that all the remaining molecules have been indeed synthesized by the Suzuki coupling, but the stringent procedure used here assures that even if the data contains a small number of errors, the general results are correct. The procedure described above identified 1586 unique BAs that are commonly used in synthesis of bioactive molecules and for exploration of structure activity relationships, 40 the most common are shown in Figure 2. The whole set, together with calculated descriptors, may be downloaded as Supporting Information.

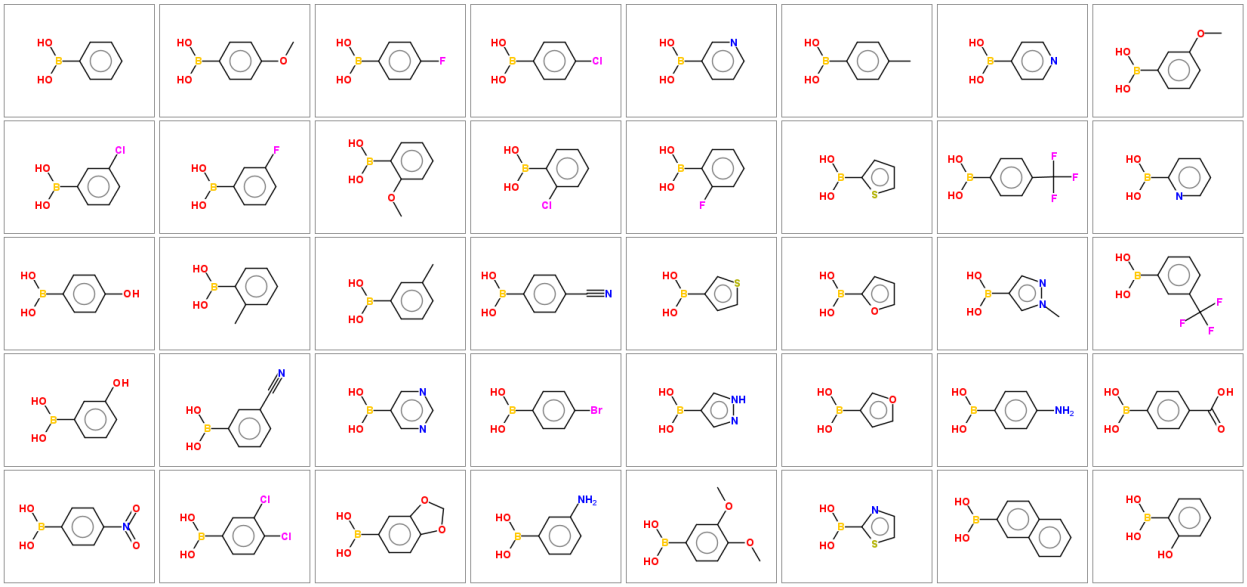


Figure 2. Boronic acids used most frequently for the synthesis of bioactive molecules (see the text for details).

2.2. Calculation of properties of boronic acids

The BAs identified as described in the previous section were characterized by 2 important calculated properties of their R-groups (the part connected to the -$B(OH)_2$ warhead) - the hydrophobicity and the electron donating/accepting power. The hydrophobicity was characterized by the calculated octanol-water partition coefficient obtained as a sum of atomic contributions similar to the scheme suggested by Wildman and Crippen.[29] Since the hydrophobicity was calculated for the R-group, not the whole molecule, this value is actually the Hansch π hydrophobicity substituent parameter. The electron donating/accepting power of the R-group was calculated using quantum-chemical calculations. Since the method is fully described in[30] and also

available online as a web service[31], only a brief description is provided here. A phenyl substituent was attached to the R-group, the resulting structure was fully optimized and the atomic charges on the phenyl calculated by the xtb 6.4.0 method considering water environment.[32] These charges are of course affected by the donating/accepting power of the attached R-group and may be therefore used to calculate the substituent parameter by the equation optimized to reproduce experimental Hammett σ constants of 89 common substituents. These 2 parameters - Hansch π hydrophobicity and Hammett σ have been suggested by Craig as optimal descriptors to analyze diversity of substituents[33,34] and used since then as classical descriptors in numerous drug discovery projects and QSAR studies.[35] Several other useful molecular descriptors were calculated for the molecules and are included in the Supporting Information data file. They include number of atoms, topological polar surface area TPSA[36] and the count of hydrogen bond acceptors and donors as used by Lipinski in formulating his Rule of 5 [37] (basically the number of O and N atoms and OH and NH groups, respectively). These properties were calculated only for the Rgroups (not considering the -$B(OH)_2$ warhead) by the Molinspiration mib engine.[38]

3. Results and discussion

3.1. Property space of bioactive boronic acids

Analysis of various data sources used in medicinal chemistry identified 11245 unique BAs. The largest portion, more than 10 thousand, is coming from the catalogs of commercial compound vendors. 1586 of them are used frequently in the synthesis of bioactive molecules. An interesting insight into the characteristics of these bioactivity supporting BAs may be obtained by comparing their properties with those of common commercially available BAs. On Figure 3 the top 200 BAs that are used in synthesis of bioactive molecules (marked by the green points) are compared with the properties of the 300 most common commercially available BAs (not containing those BAs present in the first set). The properties used for comparison are the hydrophobicity of the R-groups (the x-axis) and their electron donating/accepting power (the y-axis). This graph is therefore nothing else than a Craig plot[33] - a scheme that has been used for visualization of structure activity relationships and selection of representative molecule subsets since the early days of rational medicinal chemistry. The plot background is color coded using the distance weighted average of group membership (bioactive / common) of the neighboring molecules. One can see clear separation between these 2 classes. While the bioactivity supporting BAs tend to occupy the lower left part of the plot, i.e. they are less lipophilic and their R-groups are electron donors, the average or inactive BAs are more hydrophobic and they have electron-accepting properties. Such separation is of course not only a consequence of properties, but rather of the chemical composition and substituent pattern of the respective molecules. While the common, inactive BAs contain many halogen, alkyl and also nitro substituents, the bioactive portion of chemical space is more rich on nitrogen-containing heterocycles and rings with amine substituents, that both can form beneficial hydrogen bond interactions with the target proteins.

This quite distinct separation between active and inactive regions of chemical space of BAs may be used when selecting building blocks for SAR exploration. Particularly in the early stage of medicinal chemistry projects when the project knowledge is limited and no structural information is available to guide the derivatization effort a classical strategy is to synthesize a set of broadly diverse derivatives or prepare small libraries, test them and use this information to direct the future synthetic effort. As we have already mentioned earlier, the Suzuki coupling is one

of the most common reactions used for such derivatization and library synthesis. With more than 10 thousand BAs available commercially, however, it is not easy to select the proper subset of building blocks that would provide maximal SAR information. The data generated in this study, particularly information shown in Figure 3, can provide help in this endeavor. Using an interactive web tool (described in the following section) we selected 30 common BAs. They have been carefully selected manually, covering well the property space, particularly its bioactive region, and various functionalities. These BAs are recommended as building blocks for SAR exploration. The molecules are shown on Figure 4 and are also marked in the data table provided as Supporting Information. The BAs selected are relatively simple with 1, maximally 2 functional groups, allowing to rationalize the influence of substituent effects on bioactivity and also covering the broad range of structural and pharmacophoric features. The molecules on Figure 4 are roughly grouped by their properties, similar to the scheme on Figure 3. We also checked their pricing in the catalogs of 2 popular compound vendors (Chemspace and Molport). Most of the selected BAs may be purchased for ~50 USD per gram, only few have the price above this level. The selected set is provided as an initial suggestion, and it may be, of course, modified or enhanced based on the concrete project requirements.

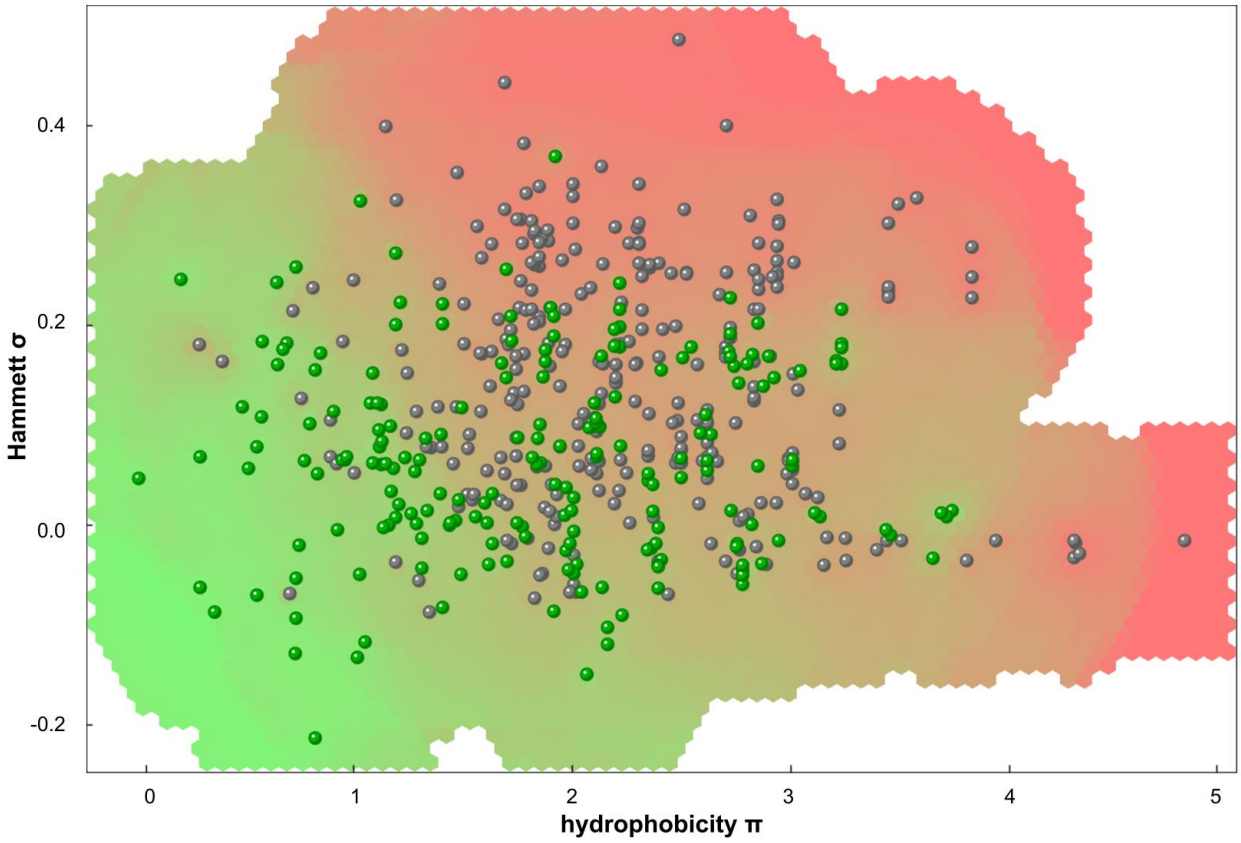


Figure 3. Comparison of properties of the most common bioactive boronic acids (green points) and commercial boronic acids (gray points). The property space beneficial for bioactivity is marked by green color.

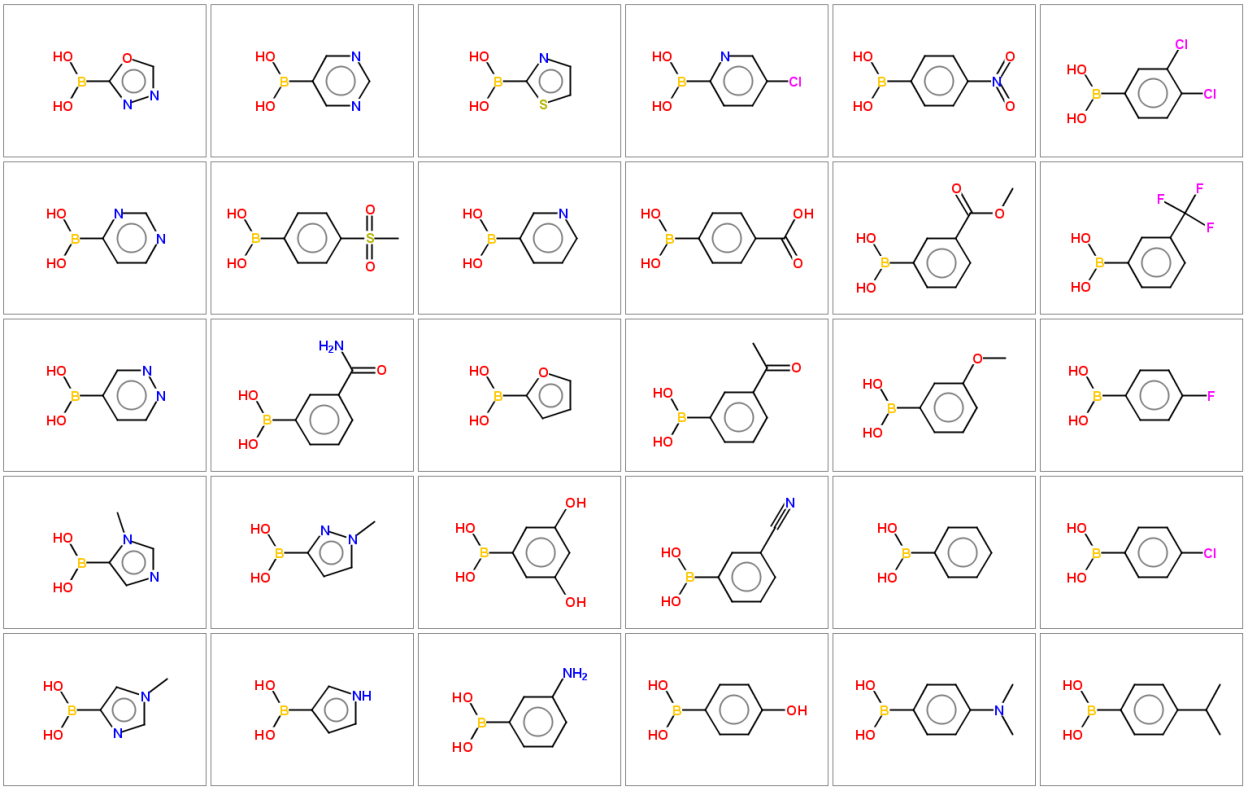


Figure 4. 30 diverse boronic acids covering well the chemical space recommended for exploring the structure-activity relationships.

3.1. Boronic acid navigator web tool

To allow interested medicinal chemists and cheminformatics scientists themselves to perform navigation in the chemical space of BAs and make their own selection based on the needs of their particular projects, a web tool offering these functionalities was developed (Figure 5). The system is written in JavaScript using the JQuery framework, the graphics is handled by the Canvas HTML element. The tool may therefore be used on standard PCs as well as on touch devices. Using this web tool the chemists are able to interactively explore the property space of 1586 BAs relevant for medicinal chemistry and make their own selection The selected molecules may be downloaded as SMILES strings also with their properties. Several keyboard shortcuts allow easy navigation within the page. For example, the 'm' key marks the 30 BAs recommended for the SAR exploration and the 't' key displays the structures for selected molecules. Additional instructions are provided directly on the tool page. The Boronic Acid Navigator is freely available at https://bit.ly/boronics.

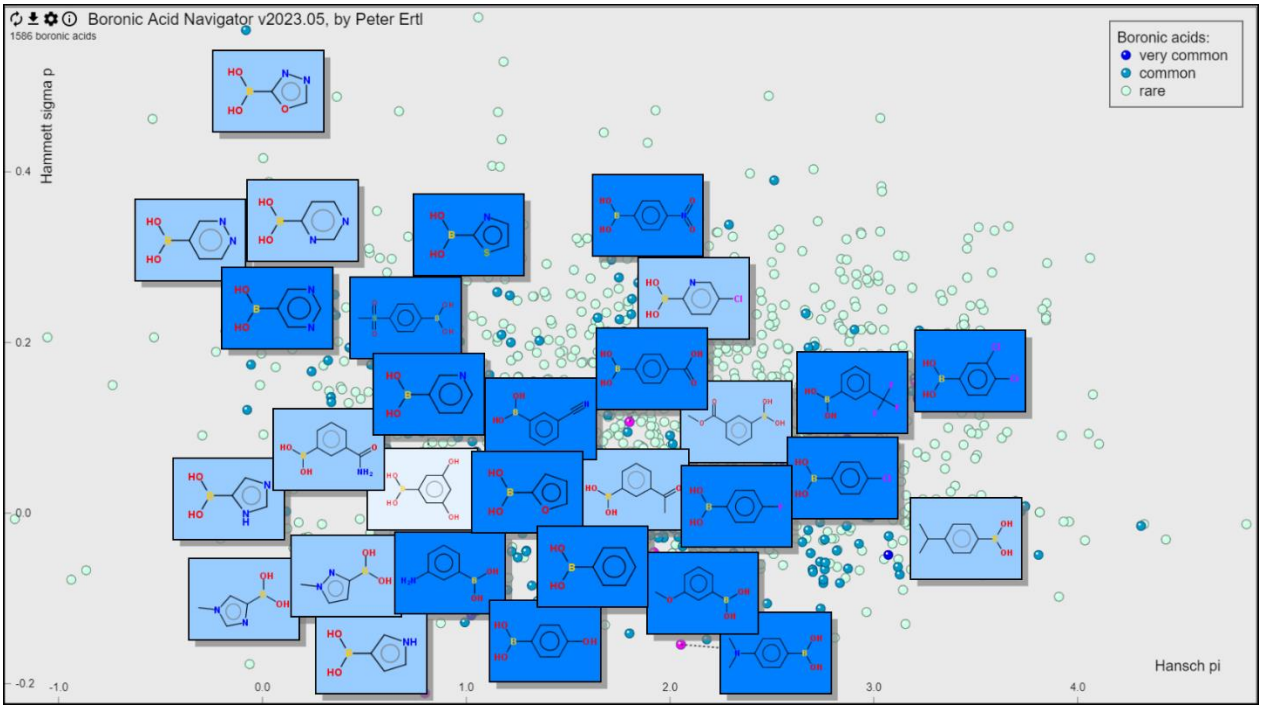


Figure 5. Boronic Acid Navigator web tool.

4. Conclusions

A large-scale analysis of 3 data sets used frequently in medicinal chemistry - catalogs of commercial compound vendors, reactions reported in patents and building blocks used in synthesis of bioactive molecules - identified 11245 unique boronic acids (BAs). These acids were characterized by relevant molecular descriptors describing their hydrophobicity and electron donating/accepting power. By analyzing property and structural space of BAs one can see a clear distinction between the acids that are often used for synthesis of bioactive molecules and the average commercially available acids. A set of 1586 BAs particularly useful in medicinal chemistry was identified and a subset of 30 well characterizing the chemical space suggested to be used in synthesis and for structure-activity exploration. A web tool - Boronic Acid Navigator - was developed that allows interactive exploration of BAs chemical space and selection of optimal set of building blocks. The web tool is available at https://bit.ly/boronics.

Data availability

All data sources used in this analysis are publicly available. Set of 1586 boronic acids used frequently in synthesis of bioactive molecules together with their calculated molecular descriptors and frequency information may be downloaded in SMILES format as Supporting Information. The Boronic Acid Navigator web tool is freely available at https://bit.ly/boronics.

Author Contributions

All authors wrote the manuscript together and all have given approval to its final version. The Boronic Acid Navigator web tool was developed by PE.

References

(1) J. Lennox, A. J.; C. Lloyd-Jones, G. Selection of Boron Reagents for Suzuki–Miyaura Coupling. *Chem. Soc. Rev.* **2014**, *43* (1), 412–443. https://doi.org/10.1039/C3CS60197H.

(2) Roughley, S. D.; Jordan, A. M. The Medicinal Chemist's Toolbox: An Analysis of Reactions Used in the Pursuit of Drug Candidates. *J. Med. Chem.* **2011**, 29.

(3) Brown, D. G.; Boström, J. Analysis of Past and Present Synthetic Methodologies on Medicinal Chemistry: Where Have All the New Reactions Gone?: Miniperspective. *J. Med. Chem.* **2016**, *59* (10), 4443–4458. https://doi.org/10.1021/acs.jmedchem.5b01409.

(4) Schneider, N.; Lowe, D. M.; Sayle, R. A.; Tarselli, M. A.; Landrum, G. A. Big Data from Pharmaceutical Patents: A Computational Analysis of Medicinal Chemists' Bread and Butter. *J. Med. Chem.* **2016**, *59* (9), 4385–4402. https://doi.org/10.1021/acs.jmedchem.6b00153.

(5) Wang, Y.; Haight, I.; Gupta, R.; Vasudevan, A. What Is in Our Kit? An Analysis of Building Blocks Used in Medicinal Chemistry Parallel Libraries. *J. Med. Chem.* **2021**. https://doi.org/10.1021/acs.jmedchem.1c01139.

(6) Monguchi, Y.; Ichikawa, T.; Yamada, T.; Sawama, Y.; Sajiki, H. Continuous-Flow Suzuki-Miyaura and Mizoroki-Heck Reactions under Microwave Heating Conditions. *Chem. Rec.* **2019**, *19* (1), 3–14. https://doi.org/10.1002/tcr.201800063.

(7) Martina, K.; Manzoli, M.; Gaudino, E. C.; Cravotto, G. Eco-Friendly Physical Activation Methods for Suzuki–Miyaura Reactions. *Catalysts* **2017**, *7* (4), 98. https://doi.org/10.3390/catal7040098.

(8) Coin, G.; Jiang, T.; Bordi, S.; Nichols, P. L.; Bode, J. W.; Wanner, B. M. Automated, Capsule-Based Suzuki–Miyaura Cross Couplings. *Org. Lett.* **2023**. https://doi.org/10.1021/acs.orglett.3c01057.

(9) Chen, J.-Q.; Li, J.-H.; Dong, Z.-B. A Review on the Latest Progress of Chan-Lam Coupling Reaction. *Adv. Synth. Catal.* **2020**, *362* (16), 3311–3331. https://doi.org/10.1002/adsc.202000495.

(10) Wu, P.; Givskov, M.; Nielsen, T. E. Reactivity and Synthetic Applications of Multicomponent Petasis Reactions. *Chem. Rev.* **2019**, *119* (20), 11245–11290. https://doi.org/10.1021/acs.chemrev.9b00214.

(11) Liebeskind, L. S.; Srogl, J. Thiol Ester−Boronic Acid Coupling. A Mechanistically Unprecedented and General Ketone Synthesis. *J. Am. Chem. Soc.* **2000**, *122* (45), 11260–11261. https://doi.org/10.1021/ja005613q.

(12) Yoshida, K.; Hayashi, T. Rhodium-Catalyzed Additions of Boronic Acids to Alkenes and Carbonyl Compounds. In *Boronic Acids*; John Wiley & Sons, Ltd, 2005; pp 171–203. https://doi.org/10.1002/3527606548.ch4.

(13) Kalliokoski, T. Price-Focused Analysis of Commercially Available Building Blocks for Combinatorial Library Synthesis. *ACS Comb. Sci.* **2015**. https://doi.org/10.1021/acscombsci.5b00063.

(14) Zabolotna, Y.; Volochnyuk, D. M.; Ryabukhin, S. V.; Horvath, D.; Gavrilenko, K. S.; Marcou, G.; Moroz, Y. S.; Oksiuta, O.; Varnek, A. A Close-up Look at the Chemical Space of Commercially Available Building Blocks for Medicinal Chemistry. *J. Chem. Inf. Model.* **2022**, *62* (9), 2171–2185. https://doi.org/10.1021/acs.jcim.1c00811.

(15) Williams, G. T.; Kedge, J. L.; Fossey, J. S. Molecular Boronic Acid-Based Saccharide Sensors. *ACS Sens.* **2021**, *6* (4), 1508–1528. https://doi.org/10.1021/acssensors.1c00462.

(16) Fujita, N.; Shinkai, S.; James, T. D. Boronic Acids in Molecular Self-Assembly. *Chem. – Asian J.* **2008**, *3* (7), 1076–1091. https://doi.org/10.1002/asia.200800069.

(17) M. António, J. P.; Russo, R.; Parente Carvalho, C.; D. Cal, P. M. S.; P. Gois, P. M. Boronic Acids as Building Blocks for the Construction of Therapeutically Useful Bioconjugates.

Chem. Soc. Rev. **2019**, *48* (13), 3513–3536. https://doi.org/10.1039/C9CS00184K.
(18) Plescia, J.; Moitessier, N. Design and Discovery of Boronic Acid Drugs. *Eur. J. Med. Chem.* **2020**, *195*, 112270. https://doi.org/10.1016/j.ejmech.2020.112270.
(19) Silva, M. P.; Saraiva, L.; Pinto, M.; Sousa, M. E. Boronic Acids and Their Derivatives in Medicinal Chemistry: Synthesis and Biological Applications. *Molecules* **2020**, *25* (18), 4323. https://doi.org/10.3390/molecules25184323.
(20) Connolly, B. A.; Sanford, D. G.; Chiluwal, A. K.; Healey, S. E.; Peters, D. E.; Dimare, M. T.; Wu, W.; Liu, Y.; Maw, H.; Zhou, Y.; Li, Y.; Jin, Z.; Sudmeier, J. L.; Lai, J. H.; Bachovchin, W. W. Dipeptide Boronic Acid Inhibitors of Dipeptidyl Peptidase IV: Determinants of Potency and in Vivo Efficacy and Safety. *J. Med. Chem.* **2008**, *51* (19), 6005–6013. https://doi.org/10.1021/jm800390n.
(21) Osakada, K.; Nishihara, Y. Transmetalation of Boronic Acids and Their Derivatives: Mechanistic Elucidation and Relevance to Catalysis. *Dalton Trans.* **2022**, *51* (3), 777–796. https://doi.org/10.1039/D1DT02986J.
(22) Wang, Y.; Xiao, J.; Suzek, T. O.; Zhang, J.; Wang, J.; Bryant, S. H. PubChem: A Public Information System for Analyzing Bioactivities of Small Molecules. *Nucleic Acids Res.* **2009**, *37* (Web Server), W623–W633. https://doi.org/10.1093/nar/gkp456.
(23) *Data Sources - PubChem*. https://pubchem.ncbi.nlm.nih.gov/sources/ (accessed 2023-04-12).
(24) Cox, P. A.; Reid, M.; Leach, A. G.; Campbell, A. D.; King, E. J.; Lloyd-Jones, G. C. Base-Catalyzed Aryl-B(OH)2 Protodeboronation Revisited: From Concerted Proton Transfer to Liberation of a Transient Aryl Anion. *J. Am. Chem. Soc.* **2017**, *139* (37), 13156–13165. https://doi.org/10.1021/jacs.7b07444.
(25) Lowe, D. *Unleashing over a million reactions into the wild – NextMove Software*. https://nextmovesoftware.com/blog/2014/02/27/unleashing-over-a-million-reactions-into-the-wild/ (accessed 2023-04-01).
(26) Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Marañón, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. ChEMBL: Towards Direct Deposition of Bioassay Data. *Nucleic Acids Res.* **2019**, *47* (D1), D930–D940. https://doi.org/10.1093/nar/gky1075.
(27) Cordero, F. M.; Giomi, D.; Lascialfari, L. Five-Membered Ring Systems. In *Progress in Heterocyclic Chemistry*; Elsevier, 2018; Vol. 30, pp 279–309. https://doi.org/10.1016/B978-0-08-102788-2.00011-8.
(28) Baumann, M.; Baxendale, I. R.; Ley, S. V.; Nikbin, N. An Overview of the Key Routes to the Best Selling 5-Membered Ring Heterocyclic Pharmaceuticals. *Beilstein J. Org. Chem.* **2011**, *7* (1), 442–495. https://doi.org/10.3762/bjoc.7.57.
(29) Wildman, S. A.; Crippen, G. M. Prediction of Physicochemical Parameters by Atomic Contributions. *J Chem Inf Comput Sci* **1999**, *39*, 868–873. https://doi.org/10.1021/ci990307l.
(30) Ertl, P. A Web Tool for Calculating Substituent Descriptors Compatible with Hammett Sigma Constants. *Chemistry–Methods* **2022**, *2* (12), e202200041. https://doi.org/10.1002/cmtd.202200041.
(31) Ertl, P. *Calculation of Substituent Properties*. https://peter-ertl.com/molecular/substituents/sigmas.html (accessed 2022-09-01).
(32) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-XTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652–1671. https://doi.org/10.1021/acs.jctc.8b01176.
(33) Craig, P. N. Interdependence between Physical Parameters and Selection of Substituent Groups for Correlation Studies. *J Med Chem* **1971**, *14*, 680–684. https://doi.org/10.1021/jm00290a004.

(34) Ertl, P. Craig Plot 2.0: An Interactive Navigation in the Substituent Bioisosteric Space. *J. Cheminformatics* **2020**, *12* (1), 8. https://doi.org/10.1186/s13321-020-0412-1.
(35) Hansch, C. Structure of Medicinal Chemistry. *J. Med. Chem.* **1976**, *19* (1), 1–6. https://doi.org/10.1021/jm00223a001.
(36) Ertl, P.; Rohde, B.; Selzer, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *J Med Chem* **2000**, *43*, 3714–3717. https://doi.org/10.1021/jm000942e.
(37) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26. https://doi.org/10.1016/S0169-409X(00)00129-0.
(38) *Molinspiration Batch Molecule Processing, v2022.10, Molinspiration Cheminformatics, Slovensky Grob, Slovakia*. https://molinspiration.com/docu/mib/index.html (accessed 2023-06-24).