Journal of Molecular Graphics and Modelling 40 (2013) 22–29

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Investigation of the influence of molecular topology on ligand binding

Rurika Oka[a,b], Ola Engkvist[b], Hongming Chen[b,*]

[a] Department of Biology, Lund University, SE-22100 Lund, Sweden
[b] Discovery Sciences, Computational Chemistry, AstraZeneca R&D Mölndal, SE-43183 Mölndal, Sweden

ARTICLE INFO

Article history:
Accepted 13 December 2012
Available online 29 December 2012

Keywords:
Ligand efficiency (LE)
Ligand lipophilic efficiency (LLE)
Ligand-efficiency-dependent lipophilicity (LELP)
Molecular topology class

ABSTRACT

Molecular topology class has previously been put forward as a new concept of describing compound quality and it has been shown that compared to general bioactive compounds, drugs is more similar to natural products and human metabolites in terms of molecular topology class distribution, in which they are enriched with compounds having only one ring system. To further understand how the molecular topology is influencing the drug discovery process, we have investigated the compound potency of different molecular topologies in published chemical patents. Our study shows that the potency for compounds having one ring system is higher compared to compounds that have more than one ring system. Compounds with one ring system are significantly less lipophilic and smaller compared to compounds with several ring systems. Further the influence of the molecular topology on ligand efficiency (LE), ligand lipophilic efficiency (LLE) and ligand-efficiency-dependent lipophilicity (LELP) was also analyzed and it was found that on average compounds with fewer ring systems and in particular compounds with only one ring system show consistently better LE, LLE and LELP. The results suggest that compounds with fewer ring systems and in particular compounds with only one ring system have good properties and that they might be useful starting point for drug discovery projects.

1. Introduction

Since the beginning of the genomics era in the 1990s, the screening strategy in a drug discovery project has been largely transformed from low throughput *in vivo* based phenotypic screening to a target centric high-throughput *in vitro* screening cascade. This paradigm shift has so far not been able to increase the productivity in the pharmaceutical industry as a whole. Recent data indicate that only 10% of the compounds that are nominated as clinical candidates will become approved drugs [1,2]. The complexity of the drug discovery process is well recognized. A successful drug discovery project involves optimization of multiple parameters, such as potency on the primary target, selectivity *versus* other targets and absorption, distribution, metabolism, excretion and toxicity (ADMET) properties. It is a very difficult path beginning with hit identification then lead generation followed by lead optimization to achieve the optimal molecular properties needed to nominate a clinical candidate. For a molecule to reach clinical candidate nomination, it may need to have high potency on the primary target, be (in most cases) highly selective and have a good ADMET property profile as well as a good PKPD (pharmaco-kinetic and pharmaco-dynamic) profile. High potency on the primary target is an important factor to optimize during the drug development process to be able to reach an optimal PKPD profile. It contributes, in most cases, to a low required dose for achieving the desired pharmacological effect. The risk of having compound related safety liabilities decreases with a lower dose. However, other factors like ADMET properties, target engagement and residence time are very important as well. In many cases high potency and good ADMET properties are very challenging to combine in one molecule [3–6]. Finding the right balance between them is still a very challenging task for medicinal chemists.

As a new descriptor for compound quality assessment, molecular topology class has been proposed. Molecular topology class is an extension of the molecular framework concept [7,8]. Comparing to molecular framework analysis, molecular topology class analysis consists of only a few classes, which makes it possible to apply rigorous statistical methods to analyze the results in an informative way. In comparison a large dataset of molecules may consist of thousands of different molecular frameworks. Previously it has been observed that there is an enrichment of compounds with only one ring system (1TR) among drugs compared to clinical candidates and general bioactive compounds. This conclusion is still valid after correcting for any potential bias originating from lipophilicity, molecular size and differences in target distribution among the different molecular topology classes [7]. The analysis was further extended with the inclusion of natural products and human metabolites and it was found that they also were enriched with compounds having only one ring system. It is particular true for human metabolites. Further analysis showed that published

* Corresponding author.
E-mail address: hongming.chen@astrazeneca.com (H. Chen).

1093-3263/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2012.12.004

compounds are getting less similar to human metabolites over time. These results might be due to the change in medicinal chemistry practice during the last decades. Analyzing the shift in how often different topological classes occur in medicinal chemistry publications indicated that this is the case. The influence of the molecular topology on the selectivity of drugs has also been investigated [9]. Compounds with only one ring system were found to be more selective than compounds with two or more ring systems. To further understand the differences between molecular topology classes, it was decided to investigate a large set of compounds extracted from a commercial patent database covering small molecule modulators of therapeutically relevant proteins. The reason for choosing patent data is that the compounds should have been reasonably well characterized by the organization filing the patent. Patent databases are the major source for analyzing compounds synthesized and utilized in the pharmaceutical industry. It is acknowledged that an analysis of data from a patent database mixes different type of targets and assays, so the data needs to be analyzed carefully to remove this type of bias. Only dose–response data that were used in the analysis and only activity data that were associated to one human gene id were used. The patent data was used to investigate the frequency of different topological classes as a function of size and lipophilicity and to investigate the influence of topology on efficiency indices for ligand binding to proteins. The chosen efficiency indices, ligand efficiency (LE), ligand lipophilic efficiency (LLE) and ligand-efficiency-dependent lipophilicity (LELP) are commonly applied in the drug discovery process.

As it has been discussed earlier, screening of compounds purely on the potency could mean selecting compounds with poor ADMET properties [3,4]. Ligand efficiency (LE) [10] has been proposed as a measure for selecting lead compounds and is used to normalize the potency with respect to the molecular size. It is therefore possible to compare potencies of compounds in a wide range of molecular size. LE is defined as:

$$\mathrm{LE} = \frac{\Delta G}{\mathrm{HEV}} \tag{1}$$

where $\Delta G = -RT \ln \mathrm{Ki}$, Ki = XC50/2 and HEV is the number of heavy atoms.

There are criticisms that LE does not completely eliminate the size bias [5,6,11]. Despite of this, LE is still commonly used to assess the affinity of compounds for a target. Especially in fragment based drug design, LE is widely accepted and practically used, since this index gives priorities to small molecules with relatively lower potency rather than large molecules with higher potency [12,13].

Another useful ligand efficiency index is ligand lipophilicity efficiency (LLE) [14], which is defined as:

$$\mathrm{LLE} = \mathrm{pXC50} - C \log P \tag{2}$$

LLE can be used to improve the potency while keeping the lipophilicity low [14]. It was shown by Tarcsay et al. [15] that increasing potency during hit identification and lead optimization also increases lipophilicity. LLE is a useful index to control the lipophilicity while improving the potency, since high lipophilicity tends to cause adverse ADMET properties and high promiscuity [14]. It has been shown that LLE can be used to differentiate between hits, leads, and drugs [15].

LE and LLE highlights low potency compounds with small size and low lipophilicity, which might otherwise be overlooked when analyzing hits from a high-throughput screening. The use of a small and less lipophilic molecule as a starting point is beneficial as the size and lipophilicity generally will be increased during the lead optimization process [4,16]. As LE does not consider the lipophilicity and LLE does not consider the molecular size, ligand-efficiency-dependent lipophilicity index (LELP) has recently been proposed to address this issue [16]. LELP is a descriptor which combines the three important factors, lipophilicity, molecular size, and potency into one descriptor. LELP is calculated using the following formula:

$$\mathrm{LELP} = \frac{C \log P}{\mathrm{LE}} \tag{3}$$

It has been observed that LELP can differentiate between drugs from hits, leads and even from compounds that had entered phase II clinical trials, which was not possible with LLE [15]. It has also been suggested that the combination of molecular mass and lipophilicity provides a better prioritization of compounds during the screening process [3]. While LELP has not been used as widely in publications as LE and LLE, it was still felt that it would be useful to include a measure that combines both size and lipophilicity in the analysis. Also, other compound efficiency metrics like size-independent ligand efficiency (SILE) [6] were investigated; the results were found to be consistent with that of LE, LLE and LELP and therefore they will not be further discussed here.

Thus LE, LLE and LELP have been shown to be useful indices to judge compound quality. In the current study, the influence of the molecular topology on LE, LLE and LELP for a set of compounds from the patent literature is investigated. The aim is to better understand

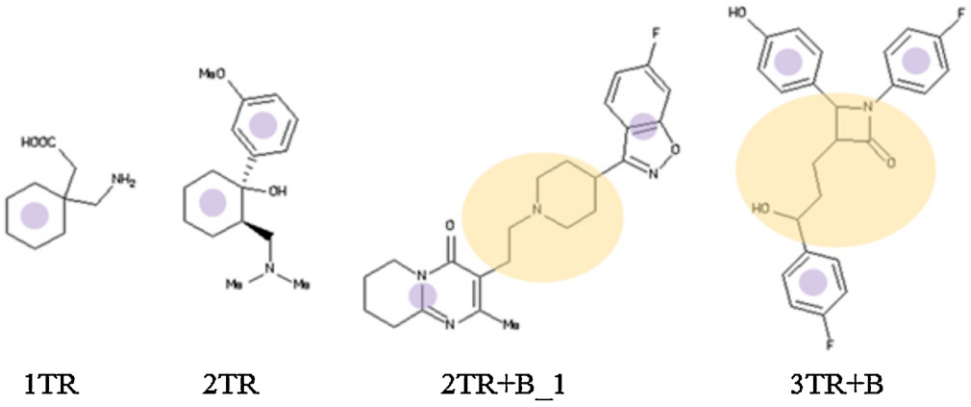


Fig. 1. Examples of compounds from different topological classes (TR = terminal ring(s), B = bridge). Purple is showing the terminal rings and orange is showing the molecular bridge in each compound. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

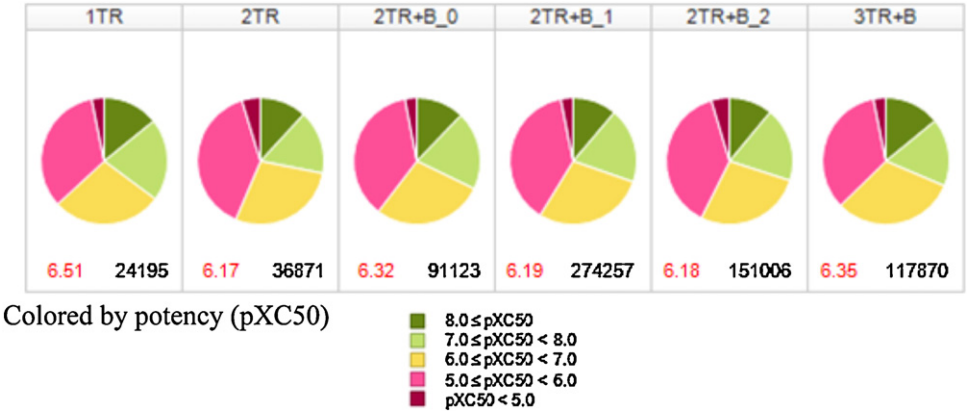


Fig. 2. The potency distribution for the six topological classes. The number of molecules represented by each pie-chart is shown in black and the average pXC50 values are shown in red.

why certain molecular topologies are more frequent among drugs than among general bioactive compounds.

2. Materials and methods

2.1. Data sets

All the data used in current study was retrieved from a commercial source [17]. Compounds with a reported potency for at least one human target were extracted. In total 1,022,057 data points which correspond to bioactivity data (pXC50) for 695,322 unique structures and 972 human targets. Target classes represented in the dataset were GPCRs, Kinases, Ion Channels, Transporters, Proteases, Nuclear Hormone Receptors, Phosphatases, Oxidoreductases and Hydrolases. All compounds with a substructure composed of more than three amino acids were removed to assure there was no influence of patents covering peptides.

2.2. Property calculations

C log *P* was calculated using a commercially available program [18]. Molecular topology classes and the number of heavy atoms (*i.e.*, non-hydrogen atoms) were calculated with an in-house C++ program based on Openeye toolkit [19]. Wilcoxon rank-sum test [20] was performed with the open source statistic package, Python-statlib [21].

2.3. Definition of molecular topology classes

The molecular topology class definition [7] is an extension of the molecular framework (MF) concept [8]. A molecule is divided into three subunits: terminal ring systems (TR), a molecular bridge (B), and side chains. A terminal ring system refers to a ring system which has only one connection to other ring systems in the molecule. The molecular bridge connects all of the terminal rings. The difference between the definition of a molecular bridge and the definition of a linker is that a molecular bridge might include additional ring systems, while a linker does not include any ring system at all. Any ring system which is directly connected through linkers to more than one other ring system is regarded as part of the molecular bridge. Thus, the MF is the combination of the terminal rings and the molecular bridge. Side chains refer to atoms that do not belong to the MF. Some examples of different topological classes are shown in Fig. 1. The 2TR+B class is subdivided according to the number of ring systems in the molecular bridge to further enhance the analysis. Compounds belonging to the 2TR+B_0 class have no ring systems in the molecular bridge and the 2TR+B_1 class refers to compounds which has one ring system *etc.* In this study, 1TR, 2TR, 2TR+B (2TR+B_0, 2TR+B_1 and 2TR+B_2) and 3TR+B topology classes were considered. More complicated classes such as 4TR+B and 5TR+B were excluded due to their low occurrence in the dataset.

3. Result and discussion

The analysis is based on a set of compounds extracted from the patent literature. The distributions of the reported potencies (pXC50) for the six molecular topology classes are shown in Fig. 2.

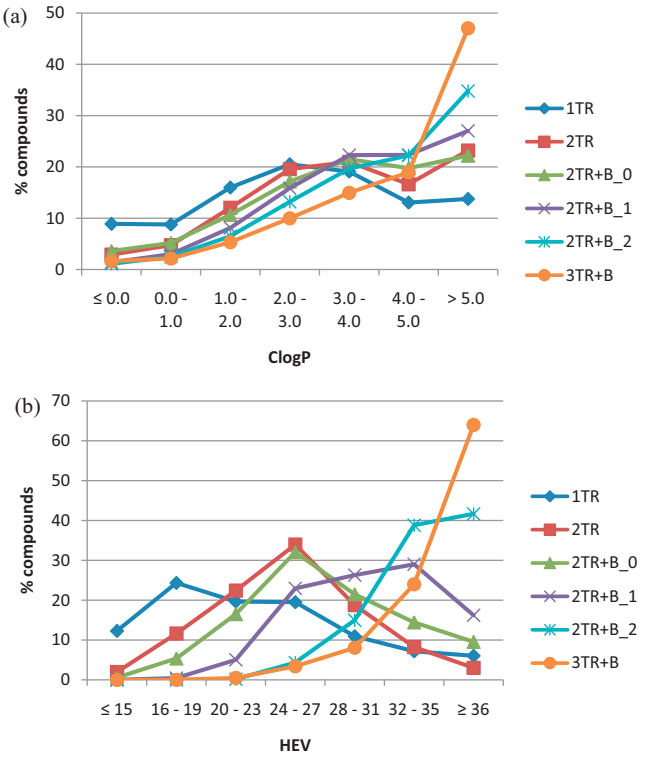


Fig. 3. The percentage of compounds in the six topological classes *versus* (a) *C* log *P* and (b) heavy atom count (HEV).

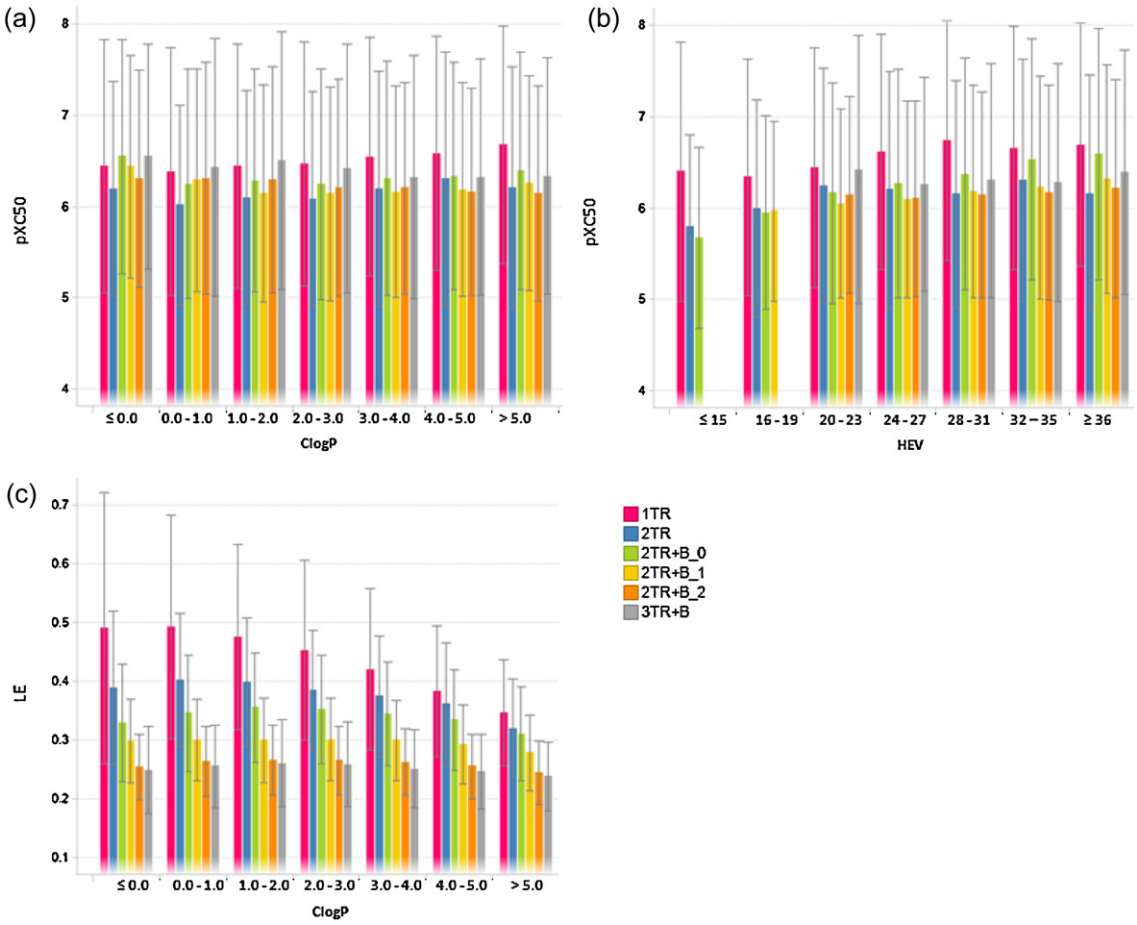
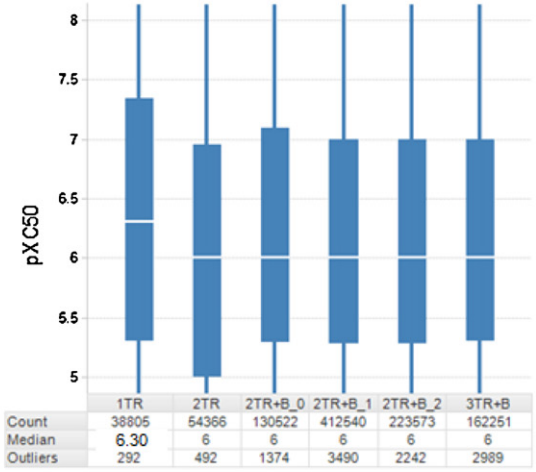
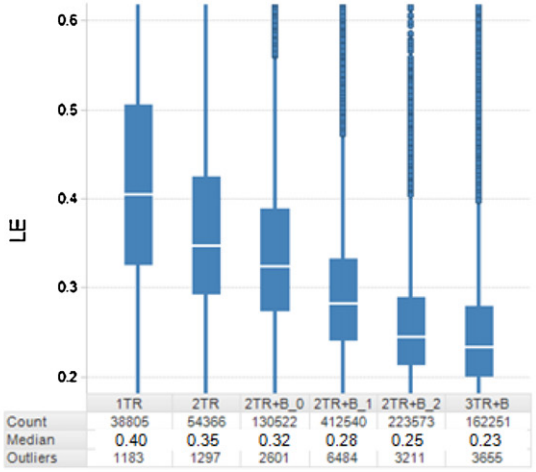


Fig. 4. The relationship between (a) *C* log *P* and mean pXC50, (b) number of heavy atoms (HEV) and mean pXC50, and (c) *C* log *P* and mean LE for the different topological classes. The standard deviations are also given in the figure.

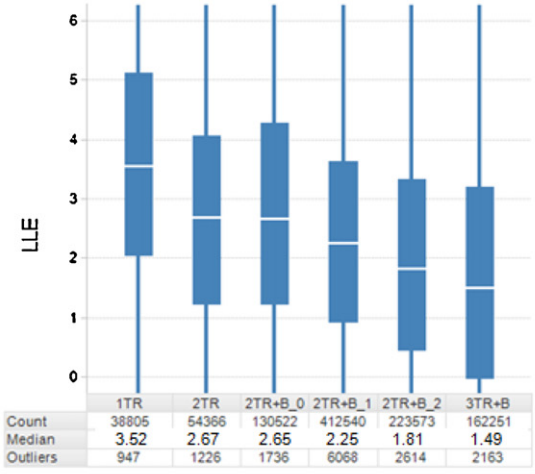
The activity distribution is fairly similar for all classes, approximately 30% of the compounds have a pXC50 ≥ 7.0, approximately 28% have a pXC50 between 7.0 and 6.0 and the rest has mainly potencies in the range between 6.0 and 5.0. Compounds belonging to the 1TR topology class have slightly higher reported averaged potency than the compounds from the other topology classes. There are also significant differences in the number of reported molecules. The class 2TR+B_1 has the most number of reported molecules, roughly 274,000 molecules, while 1TR has only 24,000 reported molecules. It reinforces the observations from an earlier study [22], which indicated a lot of compounds belonging to the 2TR+B and 3TR+B classes are synthesized in modern medicinal chemistry programs. These compounds are then identified in screening campaigns, followed-up in drug discovery projects and accordingly patented. In order to investigate the influence of the lipophilicity on the dataset, the set was binned according to *C* log *P* for each topological class. As it is seen in Fig. 3a, the distribution of the number of compounds differs a lot between the different topological classes. Compounds belonging to the 1TR topological class are more frequent among compounds with a low *C* log *P* compared to compounds from the other topological classes. Similarly it was investigated how the distribution of molecules for the different topological classes varies as a function of the size (Fig. 3b). The trend is clear; compounds with fewer ring systems are generally smaller, which is consistent with earlier investigations. Compounds with one ring system (1TR) are the smallest, while compounds with two ring system are in the middle (2TR and 2TR+B_0). This result is not a surprise since in general the molecular size increases with increasing number of ring systems. Thus, patented compounds belonging to the 1TR topology class are both less lipophilic and smaller than compounds from the other topological classes. However, the potency is not taken into account in the analysis. Thus the identified high percentage of compounds belonging to the 1TR topological class might consist of weakly active compounds, even though compounds belonging to the 1TR topology class are on average slightly more active than compounds from the other topological classes (Fig. 2). To better understand the relationship between the potency and the lipophilicity and size, the average potency as a function of lipophilicity and size is plotted in Fig. 4. For *C* log *P* values above 2.0, compounds belonging to the 1TR class are the most potent and accordingly have the highest LLE. For *C* log *P* below 2.0, compounds belonging to the 3TR+B topology class have the highest potency. Visual inspection of the compounds revealed that they are peptidomimetic compounds. However, as already shown in Fig. 3a the fraction of compounds for the 3TR+B class with a low *C* log *P* is very low. To better understand the inter-relationship between activity, size and lipophilicity, the LE was plotted *versus* *C* log *P* (Fig. 4c). It is shown that for all *C* log *P* values, the compounds belonging to the 1TR topological class has the highest LE values followed by compounds belonging to the 2TR topological class. Compounds belonging to the 3TR topological class have the lowest LE. Thus the peptidomimetic compounds belonging to the



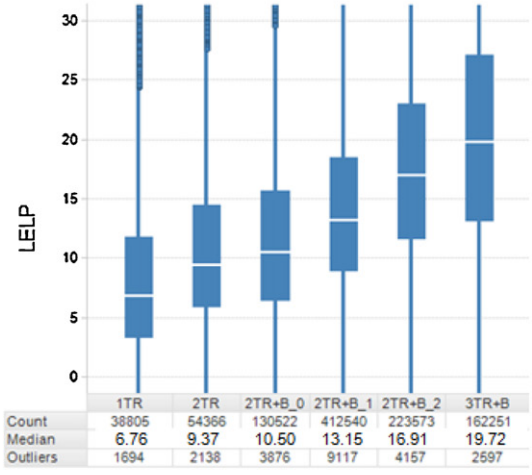
(a) Potency (pXC50)



(b) Ligand Efficiency (LE)



(c) Ligand Lipophilic Efficiency (LLE)



(d) Ligand-efficiency-dependent lipophilicity (LELP)

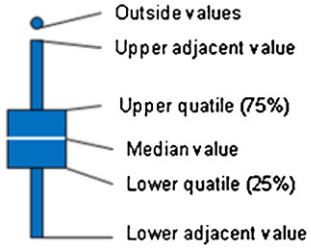


Fig. 5. The (a) pXC50, (b) LE, (c) LLE and (d) LELP distributions for each topological class. Count is the number of molecules for each topological class.

3TR+B class with a high potency and low *C* log *P* have a low LE, due to their large size. In conclusion, compounds belonging to the 1TR class have overall the best properties taking size and lipophilicity into account. More elaborated graphs describing the relationship between potency, topological class, size and lipophilicity are in the Supplementary Material (Figures S1–S4).

It is of general interest to study the impact of the molecular topology on binding efficiency indices such as LE, LLE and LELP (Eqs. (1)–(3)). The LE, LLE and LELP for the different molecular topology classes are therefore compared (Fig. 5). As already shown in Fig. 2 compounds belonging to the 1TR topology class have slightly higher potencies, however, the differences in LE, LLE and LELP are significantly larger. Besides being slightly more potent, 1TR compounds are on average significantly smaller and less lipophilic which gives rise to the differences observed for LE, LLE and LELP. The topological classes consisting of two ring systems (2TR and 2TR+B_0) have better LE, LLE and LELP than the topological classes with more than two ring systems (2TR+B_1, 2TR+B_2 and 3TR+B). The difference in LE is not only due to that the compounds belonging to the 1TR topological class is smaller, as shown in Fig. 4b, they are also more potent for a fixed molecular size.

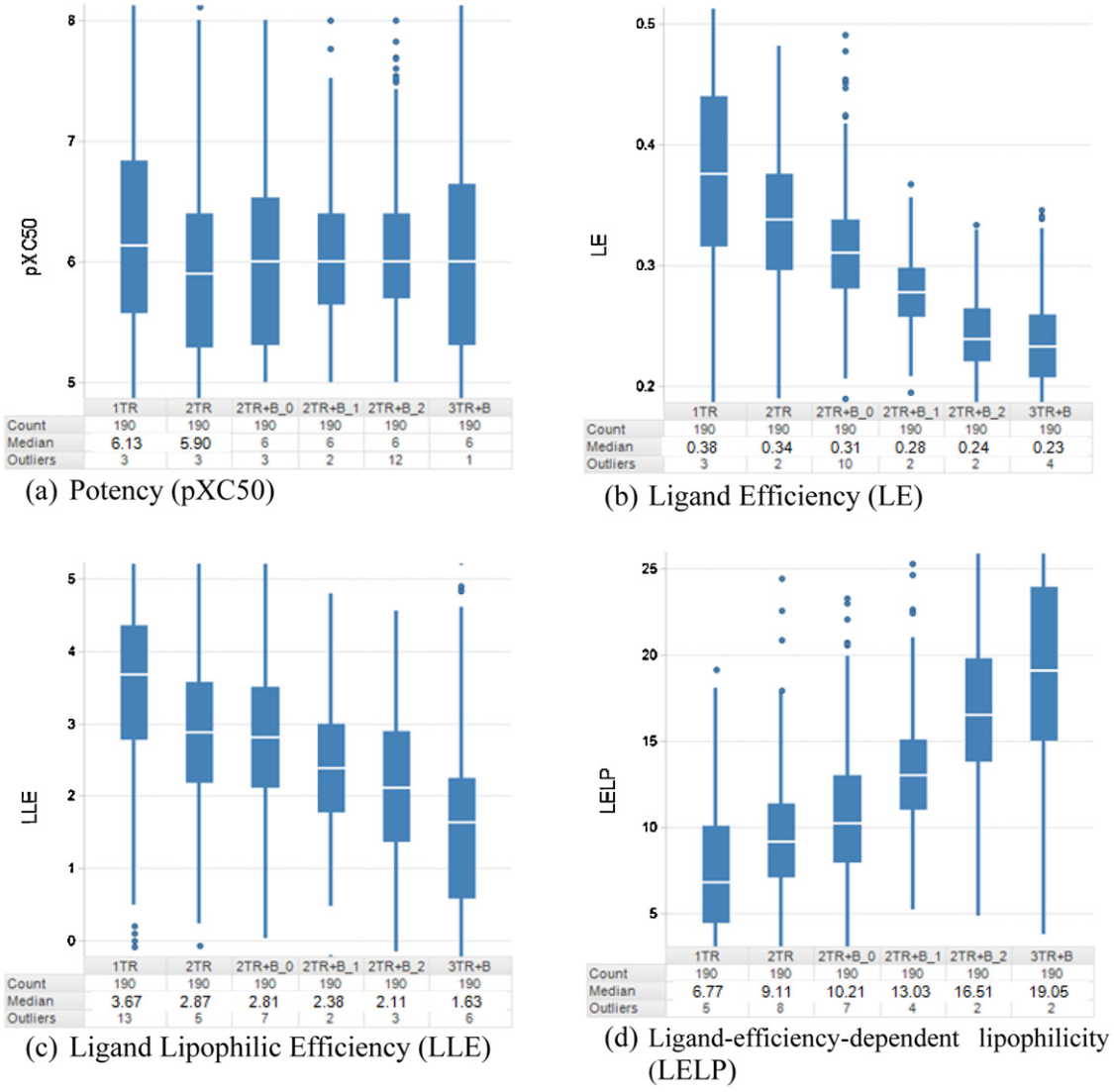
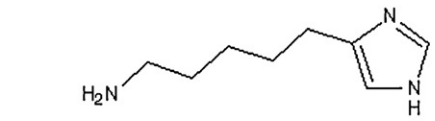
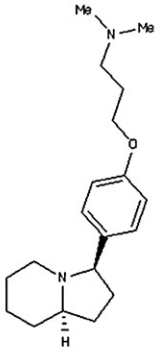


Fig. 6. The distribution of the median (a) pXC50, (b) LE, (c) LLE and (d) LELP for each target and topological class. Count is the number of targets which have more than 10 compounds for each topological class.

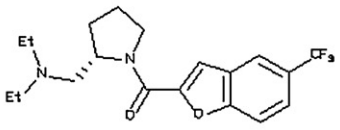
However, it should be noted that so far the analysis has not taken into account that compounds from different topological classes might be reported for different targets. In order to exclude that the identified results are affected due to the compounds belonging to different topological classes modulating different targets, a separate analysis has been done where the median LE, LLE and LELP are calculated for each target and topological class. The differences in the medians for each target are then compared for the six topological classes. Only targets where there are at least 10 compounds known for each of the six topological classes were used in the analysis. In total, 190 targets fulfilled this requirement. For each of these selected targets, the median values of LE, LLE and LELP for the six molecular topology classes were calculated (Fig. 6). Fig. 6 shows that compounds belonging to the 1TR class have the highest median values for both LE and LLE and the lowest median value for LELP. The topological classes with several ring systems have the lowest LE and LLE and highest LELP. Thus the results here are similar to the results in Fig. 5. Each pair of topological classes was also compared for all targets where at least one compound from each of the two topological classes is reported. The results are given in Table S1 and are consistent with the results in Figs. 5 and 6. The most common target classes (Enzymes, GPCRs, Ion Channels and NHRs) were also investigated individually for each pair of topological classes and the results are also consistent with the previous conclusions (Tables S2–S5). Compounds belonging to the 1TR class have on average better LE, LLE and LELP for each target class. Thus the observed trends of LE, LLE and LELP with molecular topology in Fig. 5 are still valid after correcting for any potential target bias. To get a better sense what the molecules in the analysis looks like, some structure examples are shown in Fig. 7. Here we have chosen a GPCR, Histamine receptor type 3 (H3), to illustrate what representative molecules look like. It should be noted that H3 has histamine as an endogenous agonist, which belongs to the 1TR molecular topology class; so it is not surprising that also synthetic compounds belonging to the 1TR topological class show high potency, LE and LLE for H3.



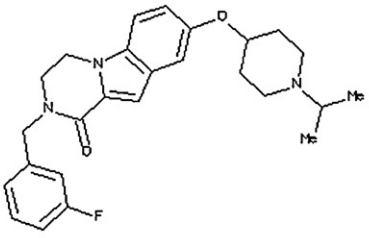
(a) 1TR
pXC50: 7.70, LE: 0.95,
LLE: 7.22, LELP: 0.49



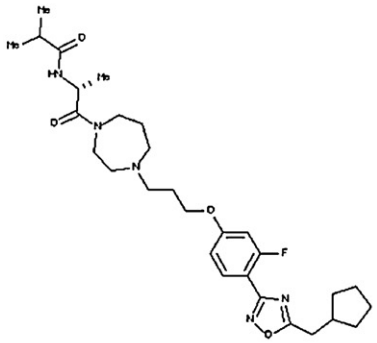
(b) 2TR
pXC50: 7.72, LE: 0.48,
LLE: 3.84, LELP: 8.11



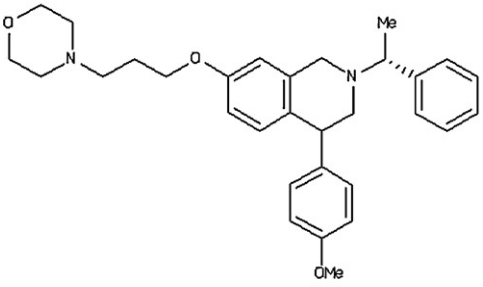
(c) 2TR+B_0
pXC50: 7.55, LE: 0.40,
LLE: 3.20, LELP: 11.0



(d) 2TR+B_1
pXC50: 7.64, LE: 0.33,
LLE: 3.03, LELP: 14.2



(e) 2TR+B_2
pXC50: 7.55, LE: 0.26,
LLE: 2.92, LELP: 17.6



(f) 3TR+B
pXC50: 7.85, LE: 0.30,
LLE: 1.70, LELP: 20.7

Fig. 7. Representative structures for the Histamine receptor H3. The median values for each topology class are: 1TR: pXC50 5.85, LE 0.52, LLE 3.67, LELP 4.35; 2TR: pXC50 5.70, LE 0.37, LLE 2.94, LELP 8.88; 2TR+B_0: pXC50 5.70, LE 0.33, LLE 2.76, LELP 8.48; 2TR+B_1: pXC50 5.70, LE 0.31, LLE 2.86, LELP 10.7; 2TR+B_2: pXC50 5.70, LE 0.24, LLE 2.65, LELP 12.3, 3TR+B: pXC50 5.70, LE 0.24, LLE 2.77, LELP 14.2.

4. Conclusions

It has been shown in earlier studies that the molecular topology influences clinical success and target selectivity. The distribution of compounds for different topological classes have also been shown to be very different for different data sources, such as drugs, clinical candidates, compounds with reported bioactivity in the literature, human metabolites and natural products [22]. To better understand how clinical success is influenced by the topological class, an analysis has been performed on compounds reported in the patent literature relevant for drug discovery. It was first observed that compounds from the 1TR and 2TR topological classes are much less reported in patents than compounds from the 2TR+B and 3TR+B topological classes. Thereafter, the distribution of the different topological classes with respect to lipophilicity and size was further analyzed and it was found that compounds belonging to the 1TR class are more likely to be smaller and less lipophilic than compounds with two ring systems (2TR and 2TR+B_0) and even more so compared to compounds with three or more ring systems (2TR+B_1, 2TR+B_2 and 3TR+B). It was also shown that at a fixed size or lipophilicity, compounds belonging to the 1TR topology class were generally more active. An exception is for the low lipophilicity range where peptidomimetic compounds belonging to the 3TR+B class were the most potent. However, only a small fraction of the compounds belonging to the 3TR+B topological class have low lipophilicity, most of them have high

lipophilicity. The analysis also showed that compounds with fewer ring systems have better LE, LLE and LELP. These trends are not due to differences in target distribution for compounds belonging to different topological classes. Thus the conclusions still hold even after taking into account that the distribution of active compounds for different molecular topology classes might be different for different targets. The reason that compounds belonging to the 1TR topological class have better LE, LLE and LELP is two-fold; first, they have on average a higher potency than compounds from the other topological classes even after taking size and lipophilicity into account. Second, compounds belonging to the 1TR topological class have generally lower lipophilicity and smaller size in comparison to compounds for other molecular topology classes. The stronger affinity for compounds belonging to the 1TR topology class might be related to the prominence of compounds belonging to the 1TR topology class among natural products and in particular human metabolites. Small molecules present in nature might have evolved to interact optimally with their target. As a consequence, synthetic molecules with a 1TR topology might also interact preferably with proteins. However, the observation that compounds belonging to the 1TR topology class have better LE and LLE is empirical. Modeling of ligand–protein interactions is still very difficult as evidenced by the difficulties to accurately calculate the free energy for protein–ligand binding. It is also an interesting observation that the percentage of compounds reported with a 1TR topology is much higher among small and hydrophilic compounds indicating that it is easier to reach a desired potency with good physico-chemical profile for compounds with fewer ring systems. However, compounds belonging to the 1TR topology class are significantly underrepresented in patent databases. This might be related to modern medicinal chemistry practice which has mainly produced compounds belonging to the 2TR+B and 3TR+B classes. These compounds are then identified as hits in high-throughput screening, followed-up in drug discovery project and thereafter reported in patents. However, the results in this article might indicate that screening more compounds similar to human metabolites might be one approach to improve the success in drug discovery projects. The human metabolite set is enriched with compounds belonging to the 1TR topological class and might therefore produce screening hits that can more easily be optimized to compounds with good LE and LLE. It should be noted that all retrospective analysis of patented compounds will reflect the historical target space that has been exploited in drug discovery projects; thus there is no guarantee that the identified conclusions will hold for future drug discovery projects that might pursue different types of targets.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2012.12.004.

References

[1] R. Peck, Driving earlier clinical attrition: if you want to find the needle, burn down the haystack. Considerations for biomarker development, Drug Discovery Today 12 (2007) 289–294.
[2] S. Paul, D. Mytelka, C. Dunwiddie, C. Persinger, B. Munos, S. Lindborg, A. Schacht, How to improve R&D productivity: the pharmaceutical industry's grand challenge, Nature Reviews Drug Discovery 9 (2010) 203–214.
[3] P. Gleeson, A. Hersey, D. Montanari, J. Overington, Probing the links between in vitro potency ADMET and physicochemical parameters, Nature Reviews Drug Discovery 10 (2011) 197–208.
[4] M. Hann, G. Keserü, Finding the sweet spot: the role of nature and nurture in medicinal chemistry, Nature Reviews Drug Discovery 11 (2012) 355–365.
[5] C. Raynolds, B. Tounge, S. Bembenek, Ligand binding efficiency: trends physical basis, and implications, Journal of Medicinal Chemistry 51 (2008) 2432–2438.
[6] S. Bembenek, B. Tounge, C. Reynolds, Ligand efficiency and fragment-based drug discovery, Drug Discovery Today 14 (2009) 278–283.
[7] H. Chen, Y. Yang, O. Engkvist, Molecular topology analysis of the differences between drugs clinical candidate compounds, and bioactive molecules, Journal of Chemical Information and Modeling 50 (2010) 2141–2150.
[8] G. Bemis, M. Murcko, The properties of known drugs. 1. Molecular frameworks, Journal of Medicinal Chemistry 39 (1996) 2887–2893.
[9] Y. Yang, O. Engkvist, A. Llinàs, H. Chen, Beyond size ionization state and lipophilicity: influence of molecular topology on ADMET properties for drug-like compounds, Journal of Medicinal Chemistry 55 (2012) 3667–3677.
[10] A. Hopkins, C. Groom, A. Alex, Ligand efficiency: a useful metric for lead selection, Drug Discovery Today 9 (2004) 430–431.
[11] W. Nissink, Simple size-independent measure of ligand efficiency, Journal of Chemical Information and Modeling 49 (2009) 1617–1622.
[12] R. Carr, M. Congreve, C. Murray, D. Rees, Fragment-based lead discovery: leads by design, Drug Discovery Today 10 (2005) 987–992.
[13] P. Mortenson, C. Murray, Assessing the lipophilicity of fragments and early hits, Journal of Computer-Aided Molecular Design 25 (2011) 663–667.
[14] P. Leeson, B. Springthorpe, The influence of drug-like concepts on decision-making in medicinal chemistry, Nature Reviews Drug Discovery 6 (2007) 881–890.
[15] A. Tarcsay, K. Nyíri, G. Keserü, Impact of lipophilic efficiency on compound quality, Journal of Medicinal Chemistry 55 (2012) 1252–1260.
[16] G. Keserü, G. Makara, The influence of lead discovery strategies on the properties of drug candidates, Nature Reviews Drug Discovery 8 (2009) 203–212.
[17] GVKBio Databases, GVK Biosciences Private Ltd., Hyderabad, India.
[18] BioByte ClogP version 4.3, http://www.biobyte.com (accessed 06.09.11).
[19] OEChem Toolkit 1.7.2, OpenEye Scientific Software, Inc., Santa Fe, NM.
[20] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics Bulletin 1 (1945) 80–83.
[21] Python-statlib version 1.1, http://python-statlib.googlecode.com (accessed 30.09.11).
[22] H. Chen, O. Engkvist, N. Blomberg, J. Li, A comparative analysis of the molecular topologies for drugs, clinical candidates, natural products, human metabolites and general bioactive compounds, Medicinal Chemistry Communications 3 (2012) 312–321.