Journal of Medicinal Chemistry

Perspective

pubs.acs.org/jmc

Rings in Drugs

Miniperspective

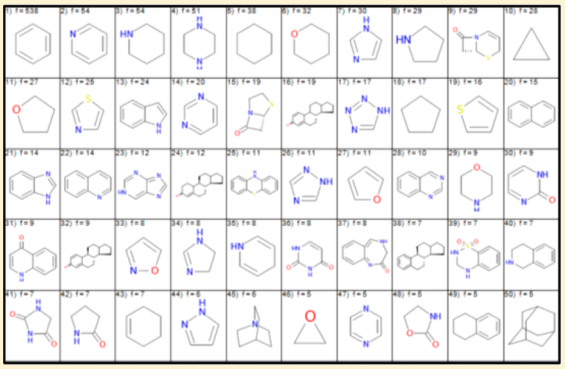
Richard D. Taylor,*[,†] Malcolm MacCoss,[‡] and Alastair D. G. Lawson[†]

[†]UCB, 216 Bath Road, Slough, SL1 3WE, U.K.

[‡]Bohicket Pharma Consulting LLC, 2556 Seabrook Island Road, Seabrook Island, South Carolina 29455, United States

Supporting Information

ABSTRACT: We have analyzed the rings, ring systems, and frameworks in drugs listed in the FDA Orange Book to understand the frequency, timelines, molecular property space, and the application of these rings in different therapeutic areas and target classes. This analysis shows that there are only 351 ring systems and 1197 frameworks in drugs that came onto the market before 2013. Furthermore, on average six new ring systems enter drug space each year and approximately 28% of new drugs contain a new ring system. Moreover, it is very unusual for a drug to contain more than one new ring system and the majority of the most frequently used ring systems (83%) were first used in drugs developed prior to 1983. These observations give insight into the chemical novelty of drugs and potentially efficient ways to assess compound libraries and develop compounds from hit identification to lead optimization and beyond.



INTRODUCTION

The concept of druglike space is ubiquitous in modern medicinal chemistry.[1,2] The application of molecular descriptors based on previously marketed drugs to assist in the design of new molecules is a widely accepted strategy to maximize the chance of clinical success.[3] Indeed the Nobel Laureate James Black stated "the most fruitful basis for the discovery of a new drug is to start with an old drug".[4] New experimental drugs have a well documented, low probability of success, with drug attrition having been attributed by Kola and Landis[5] to poor drug efficacy, toxicology, safety, and physicochemical characteristics. However, it is difficult to quantify the number of early stage projects that have not made a successful transition to human clinical trials where the failure is due to chemical issues associated with a series or a substructure, such as poor potency or undesirable absorption and distribution, metabolism, excretion, and toxicity (ADMET) profiles. In this work we are focusing on the chemistry that is associated with successful drugs and we anticipate that this analysis could be usefully applied to derisk the medicinal chemistry design process for novel therapeutic targets.

Optimal druglike parameters and characteristics are a cornerstone for medicinal chemistry, and the understanding of a successful drug profile is vital in the development of future small molecule new molecular entities (NMEs).[6] The work of Lipinski et al.[1] is commonly used when describing druglike space and is one of the most cited methods describing a likely route to clinical success based on a set of molecular descriptors for small molecule marketed oral drugs. There have been many other extensions and refinements to classify druglike space of small molecules using simple descriptors such as molecular weight, clogp, polar surface area, and number of hydrogen bond donors and acceptors.[7,8] More sophisticated methods include 2D fingerprints, 3D fingerprints, and other molecular similarity metrics.[9,10] Subsequent parallel and complementary work has focused on the leadlike property space to offer guidance for early stage molecular designs.[11]

A further extension is based on the Astex rule of three,[12] which suggests a set of molecular parameters for a fragment screening library. These knowledge-based drug design strategies typically focus on whole molecule property space such as molecular weight and hydrogen bond donors and acceptors. An alternative approach focuses on the chemical composition or substructure scaffolds of drugs[13] rather than the whole molecule 1D and 2D properties.

Key scaffold components in medicinal chemistry are the ring systems, the fundamental building blocks of most drugs on the market today. The importance of rings is well understood by modern medicinal chemists, since they play a significant role in molecular properties such as the electronic distribution, three dimensionality, and scaffold rigidity. They are often key factors in whole molecule properties such as lipophilicity or polarity and can determine molecular reactivity, metabolic stability, and toxicity.[14]

Received: November 14, 2013
Published: January 28, 2014

   dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

BACKGROUND

Paul et al.[15] showed that the combined capitalized cost per NME launch was $166 million and $414 million for hit-to-lead and lead optimization, respectively, highlighting the importance of timely and efficient choices for both the initial scaffold and subsequent development. Given the importance of chemical rings in molecular scaffolds and drug discovery, it is not surprising that there has been a range of publications on this topic. The following is a brief summary of what we consider to be some of the most important contributions in this area. They can broadly be divided into ring enumeration and the study of ring properties.

One of the earliest and most significant contributions in the area of ring enumeration was by Bemis and Murcko[16] who analyzed 5120 known drugs to identify the common ring scaffolds. They concluded that half of the drugs could be described using 32 scaffolds, where a scaffold was defined by rings and linkers ignoring atom type, hybridization, and bond order. Since this work, Lewell et al. have generated a Web based database of drug rings[17] and Ertl et al. generated a database of five- and six-membered rings[18] showing that bioactivity was sparsely distributed in relatively small bioactive islands. Siegel and Vieth analyzed the scaffold building blocks of 1386 marketed drugs[13] and found that 30% of drugs contain other drugs as building blocks. In 2010 Wang and Hou applied a systematic exhaustive approach to enumerate the scaffolds and rings in drugs[19] and showed that there was significant overlap between approved drugs and experimental drugs. They also compared their database of rings and scaffolds to the work of Bemis and Murcko, Lee and Schneider,[20] and Sigel and Vieth, noting significant differences that were attributed to the underlying databases, thus demonstrating the importance of a well curated data set for this type of analysis.

These studies have mostly focused on enumeration of druglike rings, and the individual groups have come to similar conclusions that the number of ring systems in drugs and bioactive space is currently very small and distributed in sparsely populated islands. Other groups have focused on larger data sets, for example, using compounds reported in the *Journal of Medicinal Chemistry*,[21] and part of their analysis showed that the number of new molecular frameworks per year has increased since 1959 and that these frameworks are composed of a small number of building blocks. They suggested that this increase in frameworks was due to the assembly of the same small set of building blocks in novel ways. They further postulated that this could be driven by the adoption of palladium catalyzed couplings such as Suzuki reactions.[22] Ertl[23] followed his earlier work to create a database of rings and scaffolds for scaffold hopping by analyzing ChEMBL,[24] DrugBank,[25] and ZINC.[26] Pitt et al. enumerated a list of suggested tractable ring systems that have not been synthesized.[27] Lipkus et al.[28] analyzed the Chemical Abstracts Service (CAS) registry for structural diversity and showed that a small percentage of frameworks occur in a large percentage of compounds and suggested that synthetic cost is a key factor in shaping the known chemical universe.

The properties of rings have been investigated by various groups. Gibson et al. studied the 100 most common hetereocyclic rings[29] to derive a principal components model for in vitro biological activity. Ritchie and Macdonald[30] explored the impact of aromatic ring count on compound progression and suggested that more than three aromatic rings in a molecule correlated with an increased risk of attrition in development. Young et al.[31] investigated the importance of the aromatic ring count as a molecular descriptor to assess the potential for a compound to be developed into a drug, claiming it to be the second most important descriptor after hydrophobicity.

In this study we have extended the work of both enumeration and evaluation of ring space through a historical analysis of rings, ring systems, and frameworks in the U.S. Food and Drug Administration (FDA) approved drugs. This includes not only the associated timelines for each ring system but also the property space of the rings and the application of rings in different therapeutic areas and target classes. The chemical novelty in drugs has been analyzed to assess how many new rings are used per drug and the distribution over time as the rings enter drug space. We make the distinction here between druglike space (often measured by a whole molecule similarity score), which we define as molecules having similar properties to drugs, and drug space where the exact ring has been used in a drug. This is slightly different from the more typical analysis of marketed drugs and druglike space, which often uses whole molecule properties and similarity metrics to measure correlations. One of the principal questions addressed in this study concerns innovation versus a pragmatic design strategy within medicinal chemistry and whether this correlates with an increase in success for a drug making it to market.

METHODOLOGY

The Drug List. Wang and Hou[19] noted the importance of the underlying database for any systematic analysis of molecules and their individual scaffolds and how this can have an impact on the final results. With this in mind we have used the current drugs on the market from the FDA Orange Book[32] for NMEs until the end of 2012 as our basic resource. These drugs were then cross-referenced against ChEMBL,[24] DrugBank,[25] Wikipedia,[33] *Nature Drug Reviews*,[34] the FDA Web site[32] and the *Annual Reports in Medicinal Chemistry*.[35] Oligopeptides, long chain polymers, proteins, and antibodies were removed from the data set, and for obvious reasons drugs that contain no rings, such as the antiepileptic drug vigabatrin, were ignored. Large macrocycles were considered to be a potentially special case, and we have decided to restrict the maximum single ring size recorded to be less than nine atoms. In silico preparation steps of the drugs included removing salts, standardizing charges at pH 7, and separating the individual components from combination therapies and storing the year of approval. After merging based on the different names of drugs and their active components followed by merging based on molecular structure and applying the filters previously described, there are 1175 drugs.

The exact dates in the FDA Orange Book are only given for drugs introduced after 1983, and so our timeline plots are shown from 1983 to 2012, although in certain instances compounds were compared between 1983 and 2012 and pre-1983. For this analysis the first recorded date a compound came onto the market could have been as a monotherapy or combination therapy, as we are more concerned with the first time a chemotype is used in whatever form. There is a possibility that the date of approval could be earlier if we considered registration all over the world; however, we viewed the FDA Orange Book to be the most accurate and well curated data set.

Definition of Rings, Ring Systems, and Frameworks. To analyze the drugs, a methodology was chosen, which is based on the way modern medicinal chemists will often define a molecular series, and that is based on the core "rings", "ring systems", and "frameworks". Examples of our definitions of rings, rings systems, and frameworks are shown in Figure 1 for the drugs

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

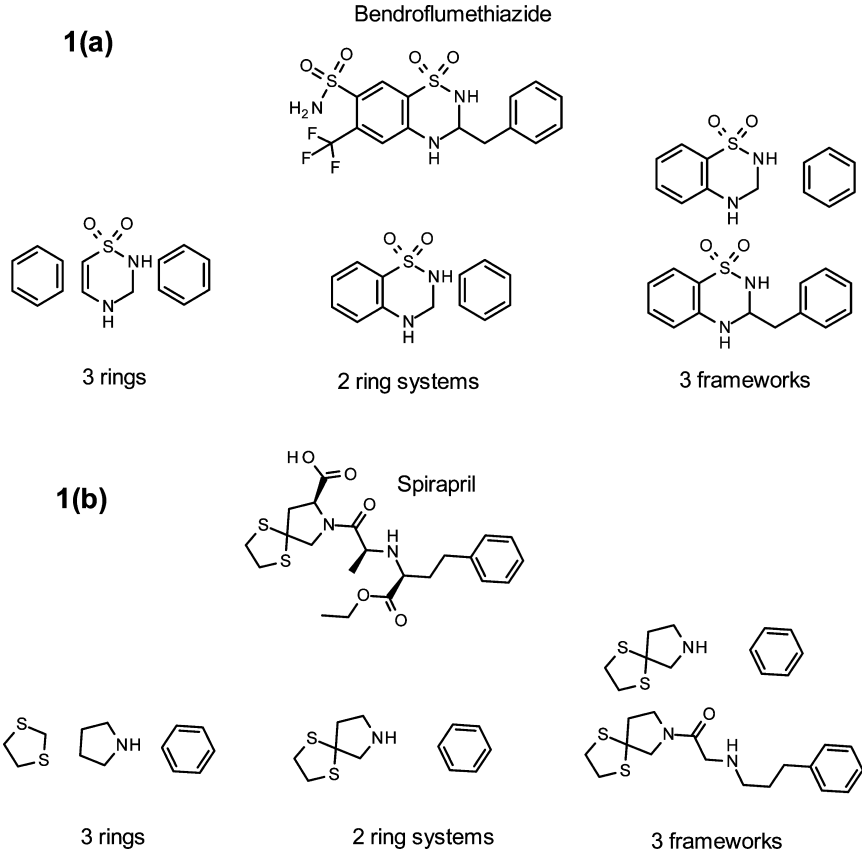


Figure 1. Example of rings, ring systems, and frameworks for (a) bendroflumethiazide and (b) spirapril.

1) All terminal non-ring bonds are removed and the framework is stored

2) A non-ring single bond is broken

3) If there are no terminal non-ring bonds and at least one ring system store the compound

4) Repeat steps 1-3 recursively until the system does not contain any non-ring bonds

Figure 2. Simplified flow algorithm for recording ring systems and frameworks. Terminal groups are non-ring groups that do no connect ring systems. Non-ring bonds exclude exocyclic carbonyls, sulfonyls, sulfinyls, thiocarbonyls, and imines.

bendroflumethiazide and spirapril. A ring is the smallest nonfused system with no acyclic (either hydrocarbon and/or heteroatom containing) linkers or terminal groups. A ring system is a complete ring or rings formed by removing all terminal and acyclic linking groups without breaking any ring bonds. In the particular case of two rings linked by an acyclic double bond, for example, the serotonin antagonist cyproheptadine, this is also included as a complete ring system and the double bond is not broken. A framework contains all the ring systems but also includes ring systems that are linked by nonterminal acyclic groups. In this analysis a distinction is made between ring systems and frameworks where a ring system can only contain ring bonds (including spiro groups) and no chain single bonds, although a framework can contain acyclic linking bonds that are nonterminating. A modified approach to the original work by Bemis and Murcko[16] was used, which retained not only the ring systems, atom types, and bond orders but also exocyclic carbonyls, sulfonyls, imines, sulfinyls, and thiocarbonyls which we believe are fundamental components of the rings.

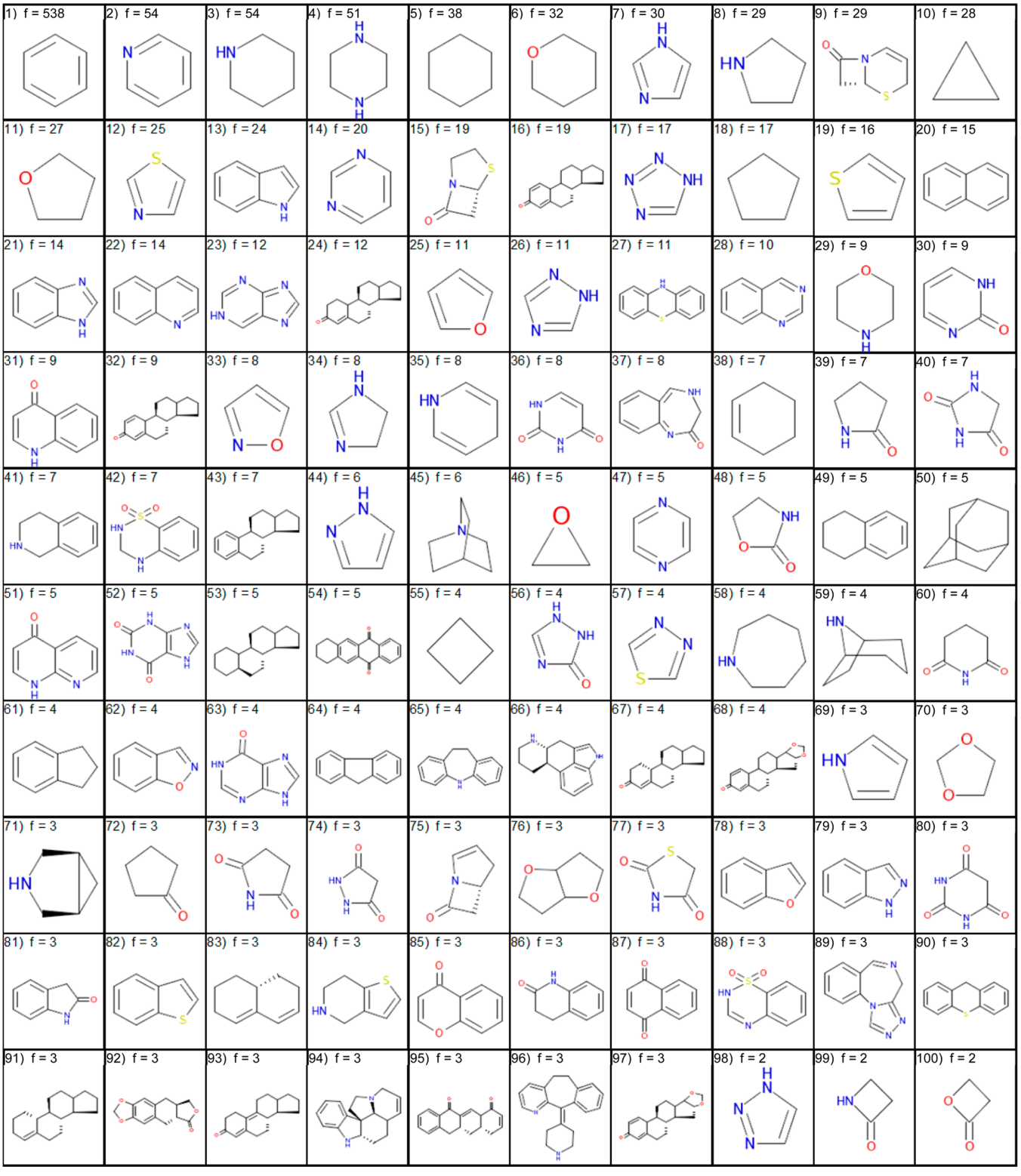
Algorithm Description. The substructure recursive algorithm was implemented in Pipeline Pilot from Accelrys,[36] and the ring systems and frameworks were recorded (see Figure 2).

We did not record the rings as a molecular database in the same way we did for the ring systems and frameworks. A ring, as defined in this work, can be created by breaking ring bonds (see Figure 1); for example, a single indole ring would be deconstructed into benzene and pyrrole. We believe breaking the ring bonds can destroy the fundamental core of the ring, and so only the number of rings was recorded for each drug because this is a property often used when analyzing molecular data sets. However, the exact molecular structures were recorded for the ring systems and frameworks as well as the overall frequency.

The cumulative frequency of each ring system and framework was stored as well as the date of approval. When the frequency of ring systems and frameworks was accumulated, all tautomers were enumerated to ensure that the exact structure matching was correct. This methodology gave separate databases for the ring systems and frameworks, and the majority of the analysis was

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

Table 1. Top 100 Most Frequently Used Ring Systems from Small Molecule Drugs Listed in the FDA Orange Book Sorted by Descending Frequency (f) and Then Ascending Molecular Weight



focused around these sets; however, it is noted based on this description that the ring systems are a subset of the frameworks.

A potential limitation of this approach, if we are simply trying to identify new fragments, is that we do not know whether the activity is due to the rings, the ring systems, the frameworks, the terminal groups, or a combination thereof. Moreover, potentially important 3D information will be lost when removing terminal aliphatic groups. Furthermore, reactive and unstable compounds could be produced from this analysis because of the computational nature in which the compounds are deconstructed. However, the primary goal of this database is not necessarily the identification of a single ring that would be used in isolation;

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

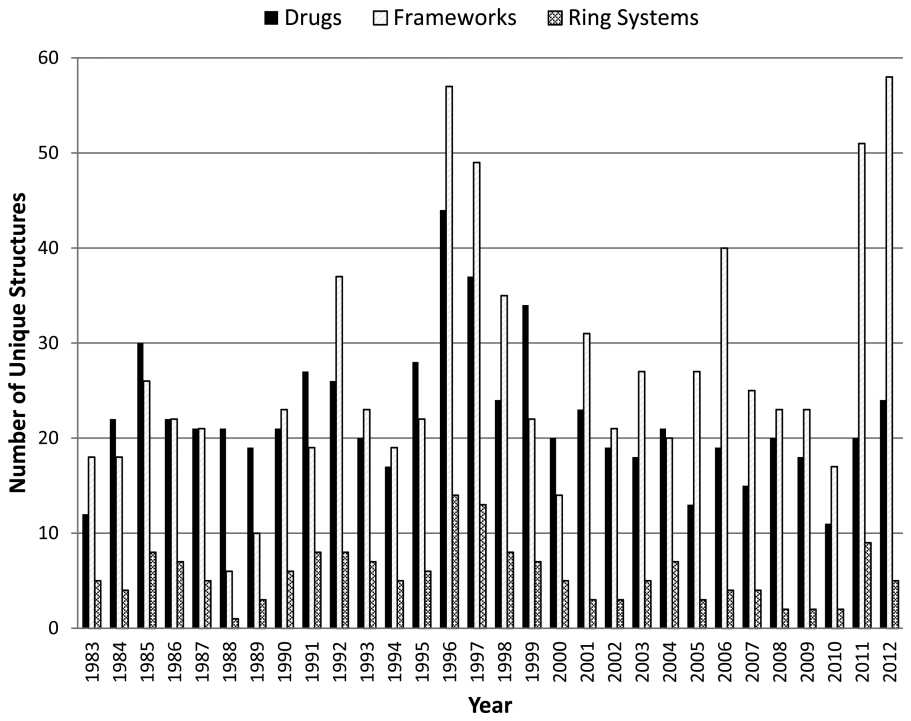


Figure 3. Total number of new drugs (small molecule NMEs using the filters previously described), frameworks, and ring systems each year from 1983 to 2012.

rather it is to generate pharmacophoric ideas or to enable substructure matching. We also take the view when comparing rings in this analysis that an exact match is the most appropriate for counting the frequency of ring systems and frameworks, although the methodology could be extended to similarity matches.

Overall, despite these limitations, we believe this is a useful approach not only to describe a set of interesting molecular scaffolds but also to show how these scaffolds are built up within a drug. The benefit of this simple approach that records ring systems and frameworks is an intuitive description of chemical space that is simple to translate to modern medicinal chemistry. Furthermore, the core ring systems are often used to describe the lead series within a drug discovery project. We also think this will give an additional level of detail when describing drug space and can be used to benchmark a library. The method can be used to assist the growth of molecular fragments and to answer the question of novelty in a molecular design process.

RESULTS

1. Ring Analysis: Total Number of Rings in Drugs and New Rings per Year. The initial analysis of the rings in small molecule drugs requires a baseline number for rings, ring systems, and frameworks in current drugs as described in the FDA Orange Book (1175 drugs). By use of the full database of 1175 drugs, there are 1197 unique frameworks and 351 unique ring systems. The top 100 ring systems are given in Table 1, and these ring systems can be downloaded as an A3 pdf poster along with the frequencies (see Supporting Information).

As previously stated, this information has been calculated in various different formats using different databases and modified descriptions of ring systems and frameworks. In this work the analysis has been extended by first looking at the associated timelines of the frameworks and ring systems. To achieve this, the total numbers of new drugs, frameworks, and ring systems were classified by the year they were first introduced to the market (see Figure 3).

The mean number of new small molecule drugs, new frameworks, and new ring systems coming onto the market each year is shown in Table 2, along with the standard deviation (sd) and median values.

Table 2. Mean, Standard Deviation (sd), and Median for the Number of New Drugs, Frameworks, and Ring Systems per Year from 1983 to 2012

	mean (sd)	median
new drugs per year	22(7)	21
new frameworks per year	27(13)	23
new ring systems per year	6(3)	5

Although the overall number of drugs has increased in recent years,[34] the number of small molecule NMEs has arguably remained constant and is generally within 2 standard deviations of 22 drugs per year, which we are referring to as the "magic 22". The only noticeable increase from the 22 per year was in 1996 and 1997. This peak has been reported by others[37] along with suggestions for the significant increase in approvals during this time, which has been attributed in part to the Prescription Drug User Fee Act (PDUFA).[38] This is discussed in more detail in the summary section. Given that there may be nonscientific reasons for this increase, we still think there may be useful lessons from this time period and we have attempted to assess whether an increase in new ring chemotypes is mirrored by an increase in clinical success. Clearly this could have implications for screening library designs and general medicinal chemistry strategies by considering new chemistry versus chemical coverage. Interestingly the peak in the number of drugs in 1996 is mirrored by an increase in the total number of frameworks and ring systems. However, in 2011 and 2012 there were 24 and 20 new small

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

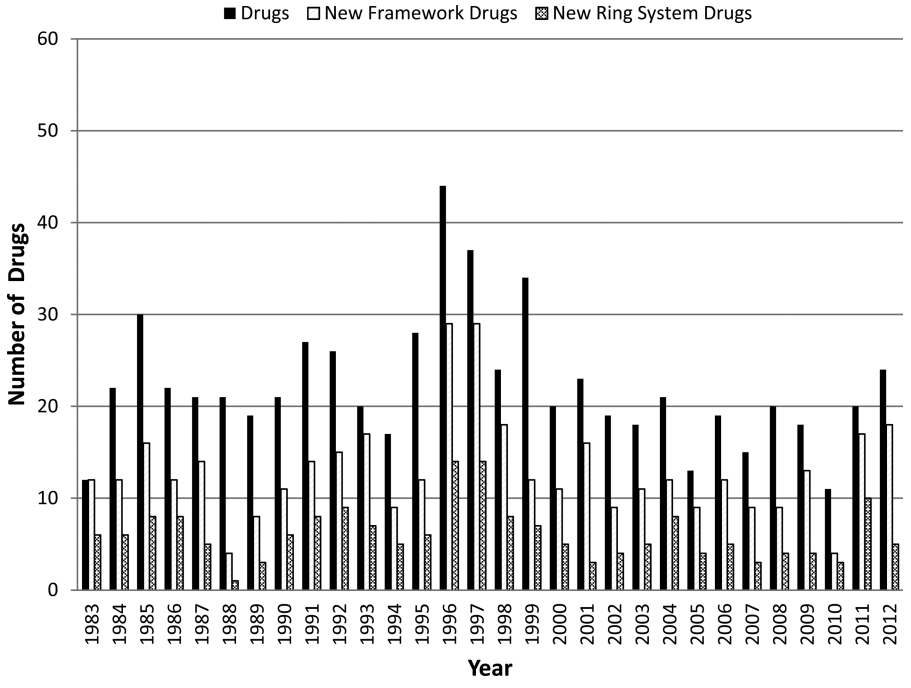


Figure 4. Total number of drugs compared with drugs containing at least one new ring framework and one new ring system.

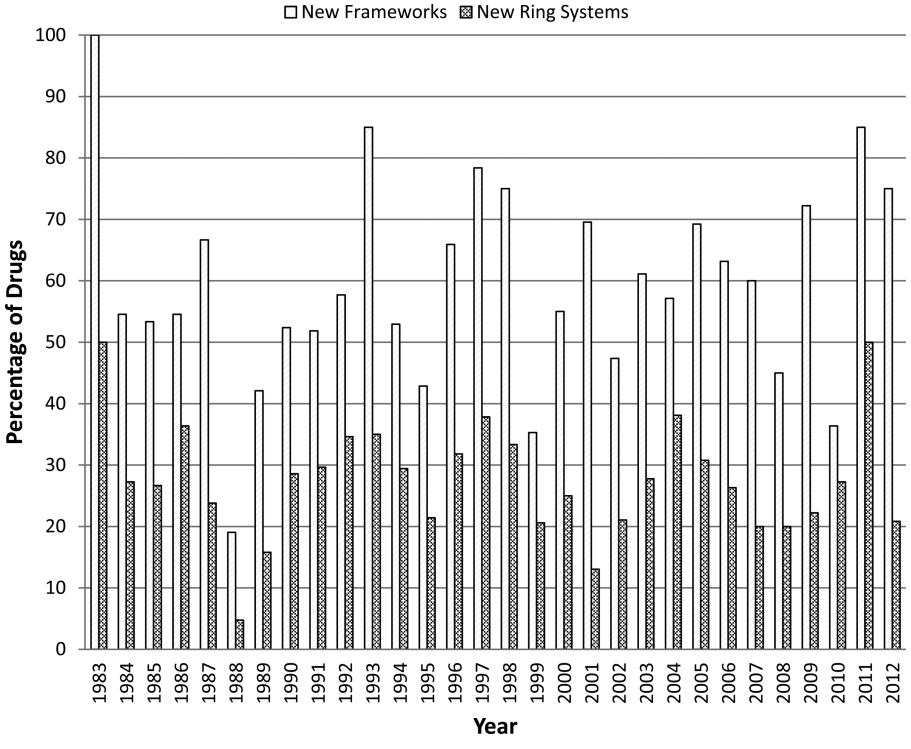


Figure 5. Percentage of drugs per year containing at least one new framework or ring system.

molecule drugs that are within 1 standard deviation of the mean, and for both years a significant increase in frameworks was observed with 2011 also showing a large number of ring systems being introduced. This could be argued as an increase in frameworks not necessarily resulting in an increase in marketed drugs, which is counter to the observation in 1996. However, a single molecule with one new central ring and many non-novel rings attached could significantly increase the framework numbers.

To overcome this issue and to investigate how the new ring systems and frameworks are distributed over time, we have looked at the total number of drugs containing at least one new ring system or framework each year rather than just the overall number of ring systems and frameworks (see Figure 4). The trends for both the ring systems and frameworks seem to match those for the complete drugs, for example, the noticeable peaks in 1996 and 1997. The data have also been plotted as a percentage

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

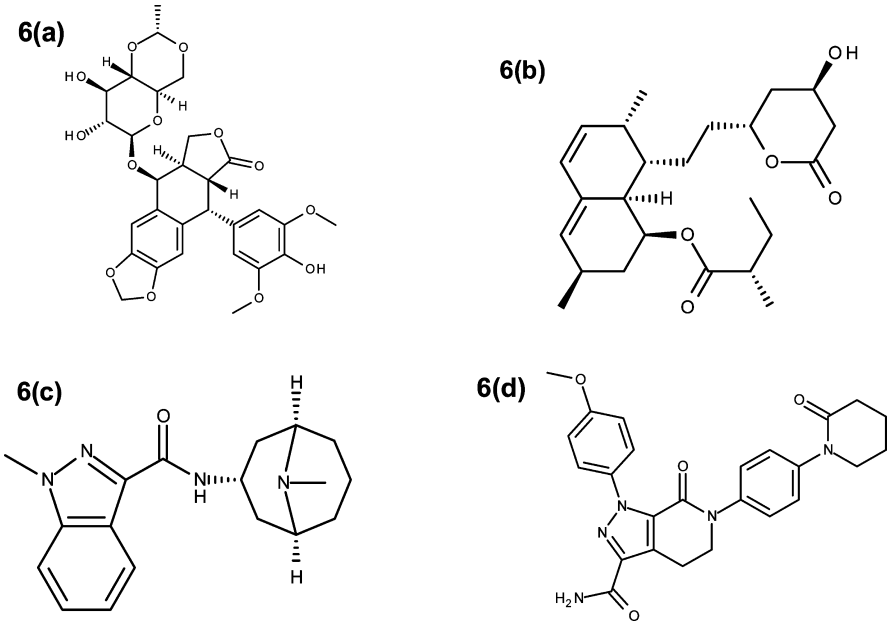


Figure 6. Examples of drugs containing more than one new ring system: (a) etoposide; (b) lovastatin; (c) granisetron; (d) apixaban.

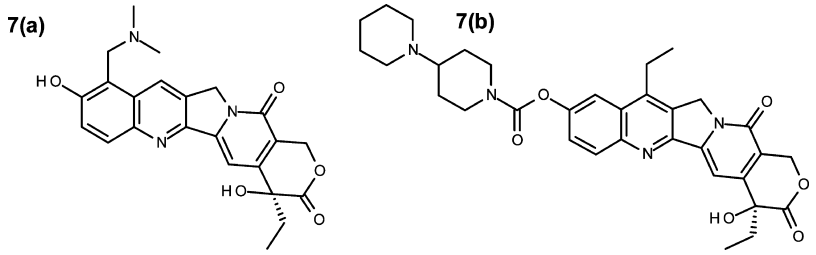


Figure 7. Examples of two drugs launched in the same year with the same new ring system: (a) topotecan; (b) irinotecan.

of drugs to normalize for an overall increase in the number of drugs (see Figure 5).

From Figure 5 the years 2011 and 1983 have the highest percentage of drugs containing novel ring systems. However, in terms of productivity 2011 is slightly below the mean value and 1983 is significantly below the average. These observations could hint at a lack of correlation between ring novelty and the number of drugs successfully making it to market. To statistically quantify this observation, we used the Kendall Tau method as a nonparametric test for statistical dependence.[39] The Tau coefficient T is given in eq 1 and is for a set of pairs of data points from (x_1, y_1), (x_2, y_2), ..., (x_n, y_n).

$$T = [(\text{number of concordant pairs}) - (\text{number of discordant pairs})]/[n(n-1)/2] \quad (1)$$

where n is the number of observation pairs. A pair of data points is considered discordant if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ and $y_i = y_j$, the pair is neither concordant nor discordant.

The Tau coefficient was calculated between the number of drugs and the percentage of drugs with new frameworks and new ring systems. If there is a perfect positive correlation between the two data sets, the coefficient is 1. Independent data sets give a coefficient of 0, and an inverse correlation gives −1.

The Tau coefficient between the number of new drugs and the percentage of drugs with a new framework was found to be −0.04 with a p-value of 0.7, which is insufficient evidence to reject the null hypothesis and leads us to conclude there is no evidence of a correlation. This calculation was repeated for the number of new drugs and the percentage of new ring systems which gave a Tau coefficient of 0.009 and a p-value of 0.9, which again indicates these data are not correlated. From this calculation we can make the following observation: There is no evidence of a correlation between the number of new drugs and the percentage with new ring systems or frameworks.

By use of the data from Figure 5, the mean of new ring systems was calculated as 28%, and as such, we were also able to make the following observation: On average 28% of new drugs each year contain one novel ring system that has not been previously used in an approved drug.

This has implications in drug discovery for both the design of a hit finding library and the optimization of an initial hit. On the basis of the two previous observations, a significant increase in the number of new ring systems in either primary screening libraries or lead optimization strategies would not necessarily lead to an associated increase in success rates for drug discovery projects. Clearly there is a need for novelty in ring systems, but a suggested alternative strategy, which has historically proved to be very successful, is to focus first on the assembly of existing drug ring systems in novel configurations.

The graphs in Figures 4 and 5 are for drugs containing at least one new ring system or framework. We could also look at the average number of new rings systems and frameworks per drug; however, it is noted that in fact there are very few drugs that contain more than one new ring system. From an analysis from 1983 to 2012 which comprises 665 molecules, only 4 drugs have

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

more than one new ring system. These drugs are the chemotherapeutic etoposide, the cholesterol-lowering drug lovastatin, the antiemetic granisetron hydrochloride, and the anticoagulant apixaban; these could be thought of as the most novel drugs with respect to the ring systems (see Figure 6). Again this could have implications on library design, as it is very unlikely that if a new ring system is used, the other rings will also be novel. This can be restated as 0.6% of new drugs contain more than one new ring system.

Also there has been one example since 1983 where two drugs have introduced the same ring system in the same year, namely, the chemotherapeutics topotecan hydrochloride and irinotecan hydrochloride in 1996 (see Figure 7). Interestingly these two drugs were successfully approved through the FDA by two different companies; GlaxoSmithKline launched topotecan hydrochloride, and Pharmacia and Upjohn launched irinotecan hydrochloride, which is now marketed by Pfizer.

2. Ring Analysis: Ring Replication and Singletons. The next step of the analysis was to assess how the ring systems and frameworks are distributed over time. We attempted to answer this question by first looking at the number of ring systems and frameworks that have been used more than once in a drug (referred to as nonsingletons) and that were first introduced prior to 1983. From the 1197 frameworks identified by this analysis, 33% were frameworks first used before 1983 and 48% of the 351 ring systems are ring systems first used prior to 1983. One could argue that this simply reflects the fact that approximately 50% of the drugs on the market were first introduced prior to 1983. However, we are more concerned with those ring systems that have been duplicated, so if we consider all nonsingleton ring systems that have been repeated after 1983, i.e., repeated in the more recent time frame, 67% of these ring systems were originally found in drugs prior to 1983. We also analyzed the top 100 most frequently occurring ring systems with the following observations: (1) 67% of the ring systems that have been repeated after 1983 were first approved in drugs prior to 1983; (2) 83% of the top 100 most frequently used rings systems are originally from drugs released prior to 1983.

The next stage of the analysis addresses how quickly a ring is reused after it first appears in a drug. Overall there are 351 ring systems. Of those, 204 (58%) have been used only once. For the remaining 147 ring systems, the time taken for the ring to be repeated has been analyzed. From those 147 ring systems that have been replicated, 65% were within 10 years. The full breakdown can be seen in Figure 8, giving rise to the following observation: 42% of ring systems in drugs have been repeated, and from those that are repeated, 65% have been within 10 years.

The singletons from the ring systems and frameworks have also been analyzed. There are 901 singleton frameworks out of 1197 frameworks, and there are 204 singleton ring systems out of the 351 ring systems. The complete timeline plot is shown in Figure 9, and the 2 years where the number of drugs have peaked (1996 and 1997) also correspond to the largest numbers of singleton ring systems and frameworks. Therefore, even though these years were productive, a lot of the scaffolds have not been reused. To further explore this observation, we have focused on the withdrawals of those drugs introduced in 1996. From this year, 2 of the 46 NMEs have been permanently discontinued and 7 of the original 46 were discontinued, although the active components have been rereleased for a variety of reasons, such as different dosing amounts. Even with these caveats around withdrawals, which are often based on nonscientific reasons, the numbers of drugs for 1996 are still significantly higher than the

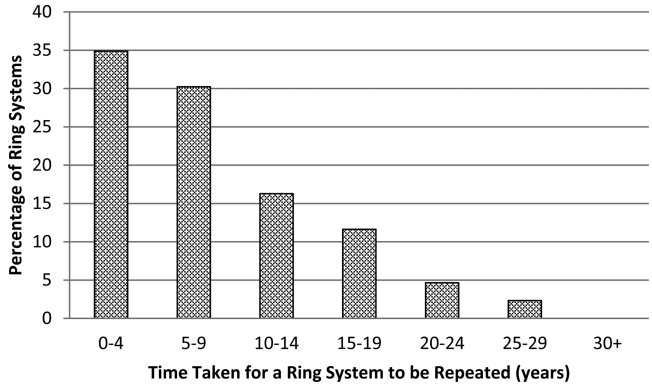


Figure 8. Time taken for a ring system to be repeated in a drug.

mean of 22. It is clear that 1996 was a productive year with a significant number of new ring systems and frameworks that have not been reused and could be a rich source of ring systems for future drugs.

Another observation is that the numbers of frameworks and ring systems are often independent. For example, in 2012 there are the largest number of frameworks and a relatively low number of ring systems, which implies that this was a year where new frameworks were produced by building up the drugs from known ring systems in a novel way.

3. Ring Analysis: Number and Size of Rings and Ring Systems. We have studied the number of new ring systems and frameworks each year and the number of new ring systems and frameworks normalized to the overall number of drugs. The next area we explored was the calculated properties of ring systems and frameworks. To do this, we first looked at the overall numbers of rings and ring systems.

The total number of rings in molecules has been widely studied,[16,17,19] and here we show the distribution of the number of rings and ring systems for all drugs (see Figure 10). We have also looked at the distribution for the subset of oral drugs (data not shown), and the distribution is very similar. We believe it is more important to look at the rings and ring systems when considering overall numbers, and it is arguably more useful to a medicinal chemist than the number of frameworks. This gives a benchmark for the complete data set and is a reminder that 95% of drugs have five rings or less and 99% have four ring systems or less. With this in mind the impact of modifying a molecular hit or lead by addition or subtraction of a ring should not be underestimated particularly if the numbers of rings or ring systems are close to these values.

The evolution over time of the number of rings and ring systems per drug has also been assessed using box plots (see Figure 11 and Figure 12) which show the quartile ranges, median, mean, and standard deviations from the period 1983−2012. We have attempted to assess whether there has been a slight increase in the mean number of rings and ring systems. A linear regression gave an increase of one ring system for both rings and ring systems over this time period, but this was a weak trend with a correlation of $R^2 = 0.45$. To see if there is any rise in the size of the drugs, we have also included a box plot for the molecular weight (see Figure 13). A linear regression on the mean showed a molecular weight increase of 50 over the same period, although the correlation was very weak with $R^2 = 0.27$, and so we do not believe this to be significant, which is also highlighted by the large standard deviations of the mean. Leeson et al.[40] have reported an increase in molecular weight over a much wider time frame and

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

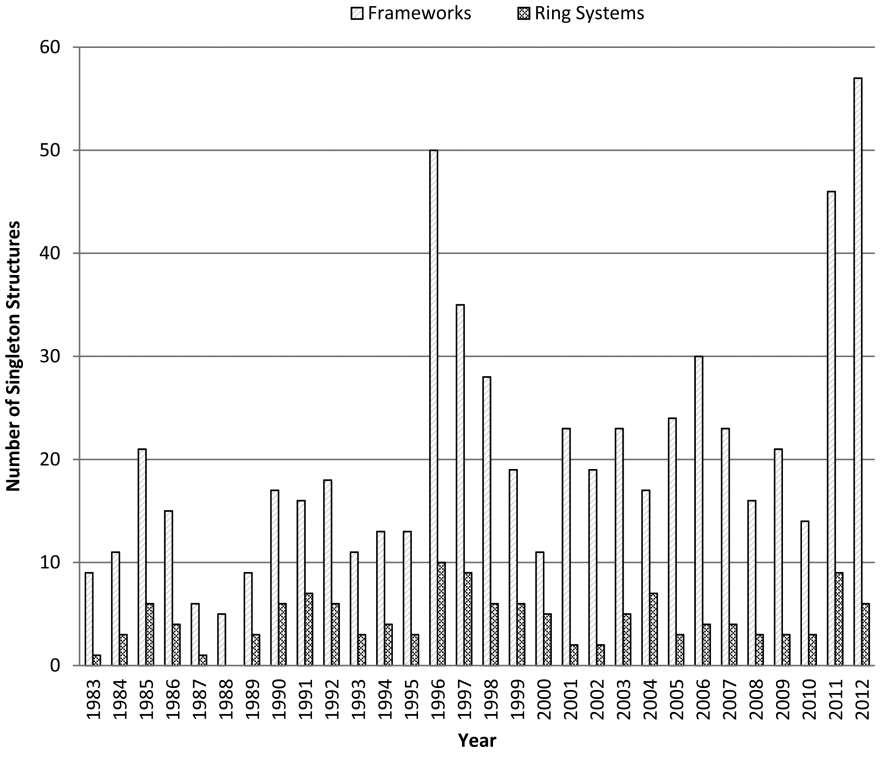


Figure 9. Number of singleton frameworks and ring systems per year from 1983 to 2012.

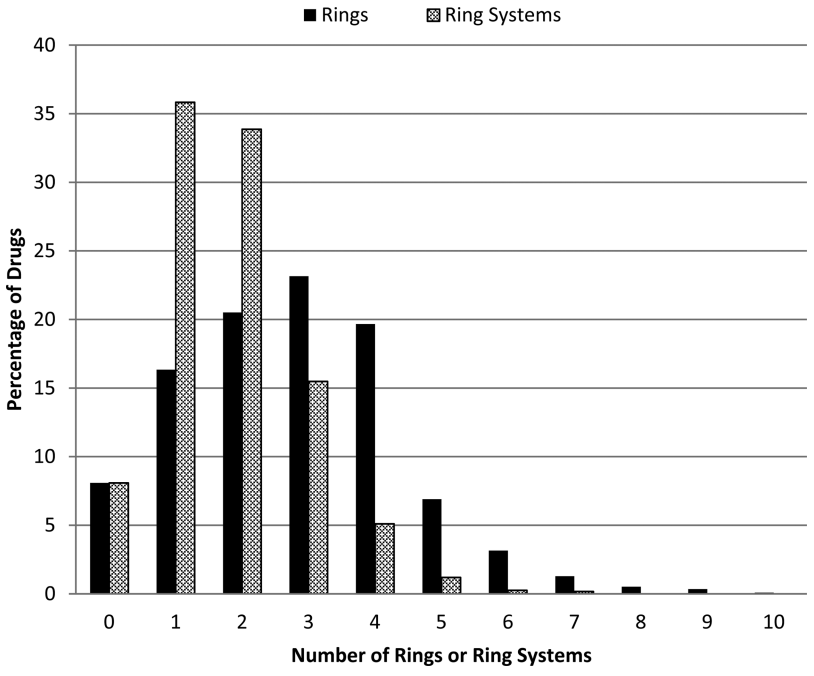


Figure 10. Number of rings and ring systems in drugs.

suggested the increase from 1930 to the present day was significant.

We have also compared the number of ring systems per drug over a longer time frame, i.e., before 1983 and 1983−2012 using mean, standard deviation, and median (see Table 3). The size of the ring systems measured by the number of atoms in the ring, excluding exocyclic groups, has also been included. An interesting observation from this analysis is that the median values for the molecular weights has increased by some 20% over this time frame, and the median number of rings and ring systems have both increased by 1. We consider this an interesting observation; however, the mean values for the rings and ring systems over the 2 time periods cannot be distinguished, since they are within 1 standard deviation.

4. Ring Analysis: Oxygen, Nitrogen, and Heteroatom Count. We have analyzed the number and frequency of a new ring system or framework entering drug space as well as the number and size of rings, but the atomic composition of those

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

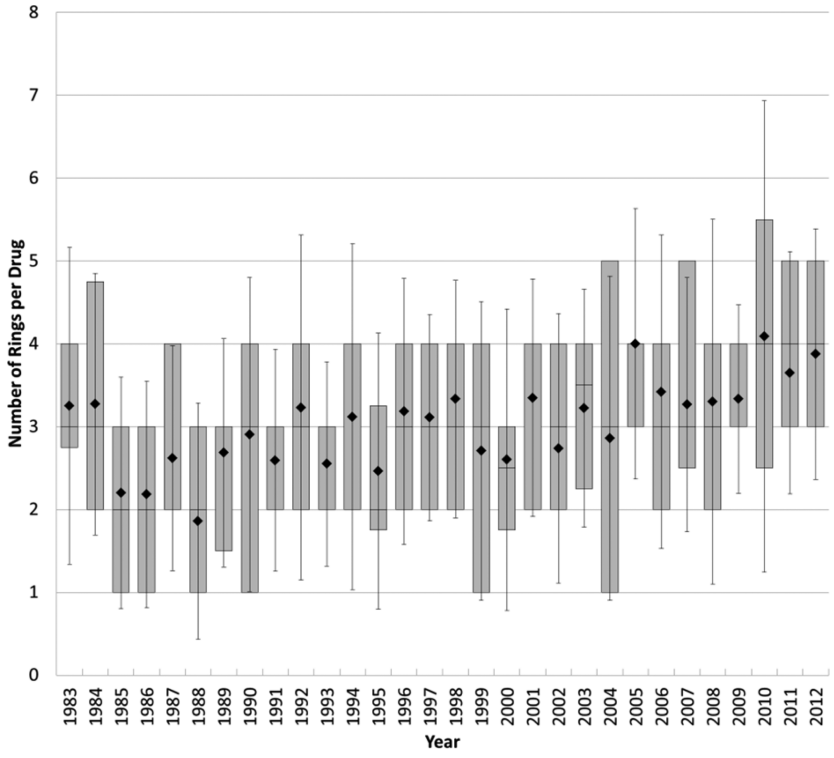


Figure 11. Box plot for the number of rings per drug from 1983 to 2012 showing the median values and upper and lower quartile ranges. The diamonds represent the mean values, and the whiskers represent the standard deviations of the mean.

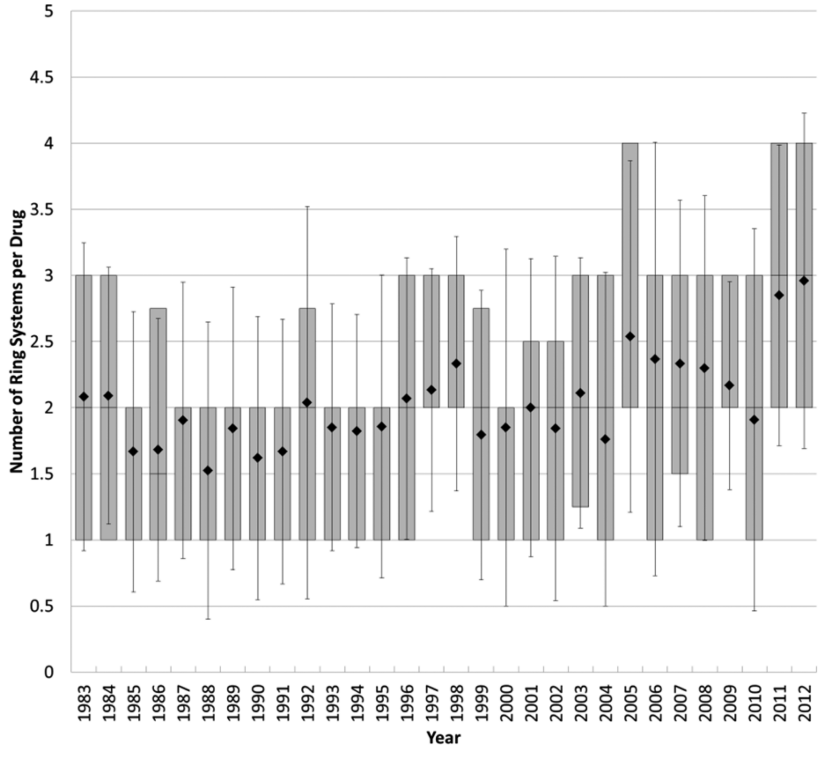


Figure 12. Box plot for the number of ring systems per drug from 1983 to 2012, showing the medians and upper and lower quartile ranges. The diamonds represent the mean values and the whiskers represent the standard deviations of the mean.

rings is also important. Owing to the nature of the algorithm used to break up compounds, we cannot record donors or acceptors. Instead, a simple and pragmatic approach was adopted to record the nitrogen, oxygen, and heteroatom (any element excluding carbon or hydrogen) count for the ring systems (see Figure 14). It is noted that for these atom counts we also include certain exocyclic groups such as carbonyls and sulfonyls as part of the ring systems. From this analysis a general conclusion can be made: The overall count of oxygen, nitrogen, and heteroatoms in a ring system is not usually greater than 3, 4, and 5, respectively.

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

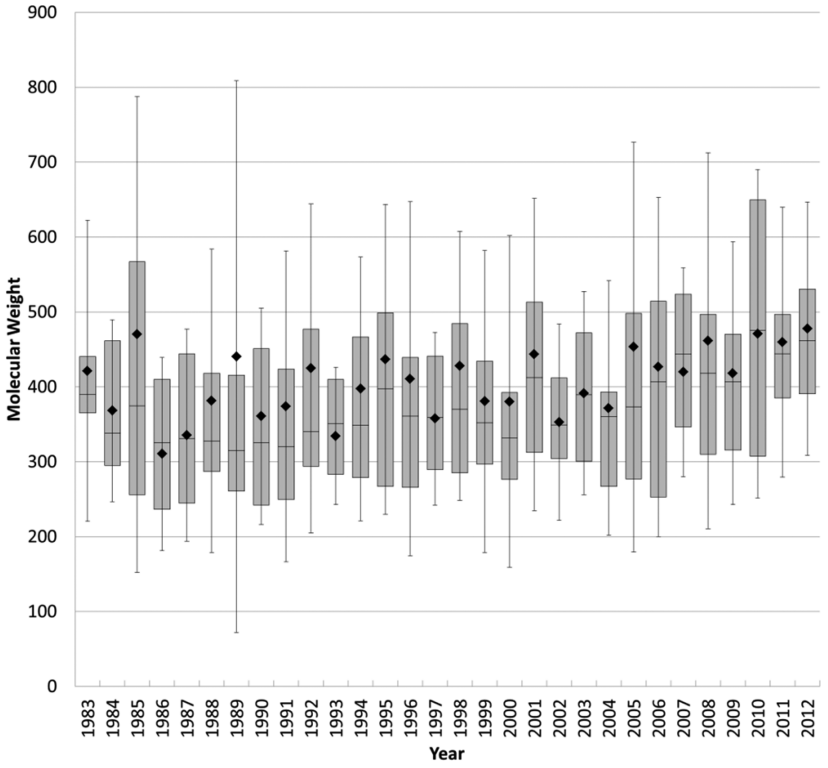


Figure 13. Box plot for the molecular weight of drugs from 1983 to 2012, showing the medians and upper and lower quartile ranges. The diamonds represent the mean values, and the whiskers represent the standard deviations of the mean.

Table 3. Comparison of Mean, Standard Deviation (sd), and Median Values for Different Properties of Drugs, Rings, and Ring Systems over 2 Time Periods

property	mean (sd), pre-1983	mean (sd), 1983–2012	median, pre-1983	median, 1983–2012
drug MW	332(168)	403(202)	302	368
no. of rings per drug	2.5(1.6)	3.0(1.7)	2	3
no. of ring systems per drug	1.5(0.9)	2.0(1.2)	1	2
no. of atoms in ring systems	11.2(5.1)	11.4(4.5)	10	10

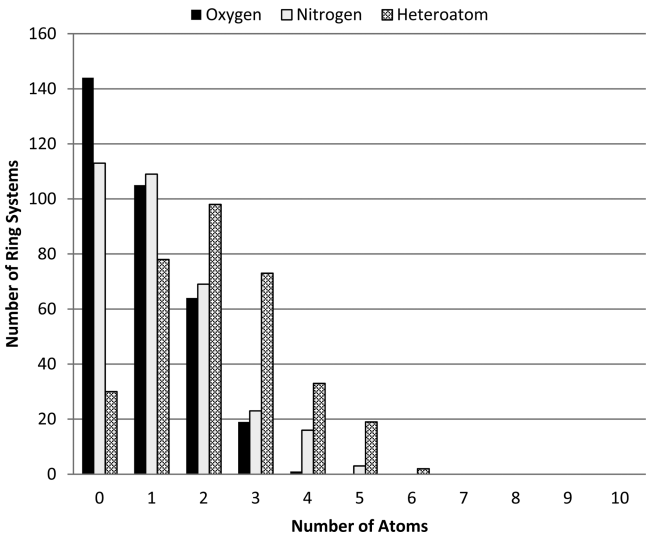


Figure 14. Oxygen, nitrogen, and heteroatom counts for all ring systems.

The next question is whether the mean atom count has changed over time. The median, mean, and standard deviations for oxygen, nitrogen, and heteroatom counts have been calculated for pre- and post-1983 (see Table 4).

The numbers for the mean and median are 1, 1, and 2 for oxygen, nitrogen, and heteroatoms for both pre- and post-1983. These data show that our use of hetereoatoms, oxygens, and nitrogens in ring systems in terms of overall numbers has not changed over time, and there is no obvious statistically relevant increase or decrease associated with these parameters.

5. Ring Analysis: Three Dimensionality and Fraction of sp^3 Carbons. Another molecular descriptor often used to assess molecular scaffolds is the fraction of sp^3 (fsp^3) carbons which is the number of sp^3 carbons divided by the number of carbons in the whole molecule (for this work all fsp^3 values are quoted as percentages). This property has been used to further investigate both the ring systems and frameworks and in particular the impact this can have on designing and assessing libraries. An example of current work in this area is by Lovering et al.[41] who suggested that there is a trend around sp^3 carbons that was typically higher for drugs compared to compounds in discovery research. However, this analysis has been questioned by Kenny et al.[42] Given this, we wanted to assess the fraction of sp^3 carbons for our database of drugs with an emphasis on the location of the sp^3 carbons, i.e., whether they are located in the rings, chain linkers, or terminal groups. Figure 15 shows the fraction of sp^3 carbons in drug molecules, as described by Lovering et al. This is

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

Table 4. Oxygen, Nitrogen, and Heteroatom Count in Drug Ring Systems for Pre-1983 and 1983−2012

property	mean (sd), pre-1983	mean (sd), 1983−2012	median, pre-1983	median, 1983−2012
oxygen count	1.0(1.0)	1.0(1.2)	1	1
nitrogen count	1.0(1.1)	1.3(1.2)	1	1
heteroatom count	2.1(1.4)	2.5(1.6)	2	2

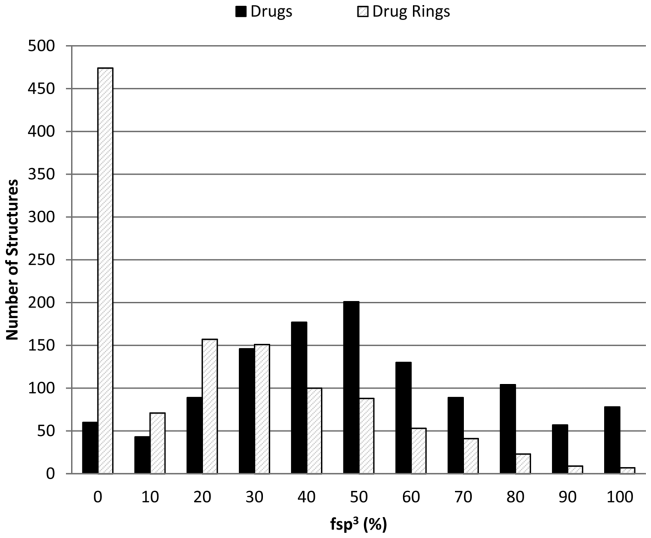


Figure 15. Fraction of sp^3 carbons and fraction of ring sp^3 carbons in drugs.

compared with the fraction of sp^3 carbons in a ring, which is simply the number of sp^3 carbon atoms that are part of a ring divided by the total number of carbon atoms in the full drug. From the histograms in Figure 15 only 5% of drugs do not contain any sp^3 carbons whereas 40% of drugs do not contain any sp^3 carbons in a ring. We believe this highlights the importance of the location of the sp^3 carbons when analyzing three dimensionality.

When the fraction is based on sp^3 carbons in rings from our database of drugs, then the following is observed: 40% of drugs do not contain sp^3 carbons in any ring or ring system. Moreover, it is widely thought that introducing three dimensionality will have a positive impact on the success of a compound.[41] Arguments could be made for increased solubility, increased options for growth vectors, and enhanced selectivity. However, these sp^3 libraries are often focused around sp^3 carbons located in the ring systems, which could incur a time penalty in slower and more challenging chemistry.[43] Given that there is a very different distribution in drugs when sp^3 carbons in rings are compared with the sp^3 carbons in the whole molecule, we have also looked at the mean distribution from 1983 to 2012 and pre-1983 (see Table 5).

The data in Table 5 demonstrate that there is no clear change over time for the percentage of sp^3 carbons in drugs (an observation also reported by Leeson et al.[40]) or ring sp^3 carbons in drugs although the median value has increased for the fraction of ring sp^3 carbons, but this is not reflected in the mean value. This is a similar observation to the atom counts in the previous section which have also been relatively consistent over these 2 time periods. We have also looked at the distribution for 1996 and 1997 where there was a spike in the overall number of drugs coming onto the market. Again this peak in the total number of drugs for these years is not mirrored in a significant change in the fraction of sp^3 carbons. A further point to note is that the fraction of ring sp^3 carbons is generally lower than those for all sp^3 carbons, although the standard deviations of the means are very high. From this analysis of ring properties it is interesting to note that the number of rings per drug, heteroatom count in rings, and sp^3 nature of rings have not significantly changed over the past 30 years.

6. Ring Analysis: Application in Therapeutic Areas and Different Target Classes. The final part of this analysis assesses the impact across therapeutic areas to understand whether the ring systems are applied to isolated targets or if they have crossed therapeutic areas or indeed target classes. In order to perform this classification, the *Annual Reports in Medicinal Chemistry*[35] was used to assign the therapeutic area for all drugs, with data from the FDA Web site to define the target class. The numbers of therapeutic areas and target classes were plotted for each ring system as well as the overall frequency, i.e., how many times the ring system had been used in any drug (see Figure 16).

The graph in Figure 16 shows that in some cases the rings have been used in only one therapeutic area and in other cases they have crossed the different therapeutic areas. The two ring systems with the highest frequency but the lowest number of associated therapeutic areas and target classes are the thiaazabicyclooctenones from the antibacterials based around cephalosporin and the anti-inflammatory corticosteroids.

From the 89 ring systems that have been used in at least two different drugs it can be observed that 63% of nonsingleton ring systems have been used in more than one therapeutic area and 72% have been used in more than one target class. This is clear evidence that ring systems can cross therapeutic areas and target classes; however, if the rings are used more than once, it is more likely it will in fact be in a different therapeutic area or target class. However, as previously stated, there are some ring systems that are important for only one target class.

It could be argued for this analysis that the difference between therapeutic area application is less important compared with target class. One might expect therapeutic areas would be agnostic to ring types and the rings in different target classes to show greater selectivity and specificity, e.g., hydroxyethylamines for proteases or aromatic donor−acceptor motifs for hinge binding in kinases.

Protein target classes can occur in multiple therapeutic areas and the class can reflect how a molecule interacts with its cognate binding site on the target protein, whereas in the case of the therapeutic areas, one might expect the same target class to be

Table 5. Fraction of sp^3 Carbons in Drugs and Fraction of Ring sp^3 Carbons in Drugs

property	mean (sd), %, pre-1983	mean (sd), %, 1983−2012	median, %, pre-1983	median, % 1983−2012
fraction of sp^3 carbons in drugs	45.3(28.1)	47.1(24.3)	42.8	46.1
fraction of ring sp^3 carbons in drugs	19.1(24.0)	20.3(21.5)	6.7	17.0

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

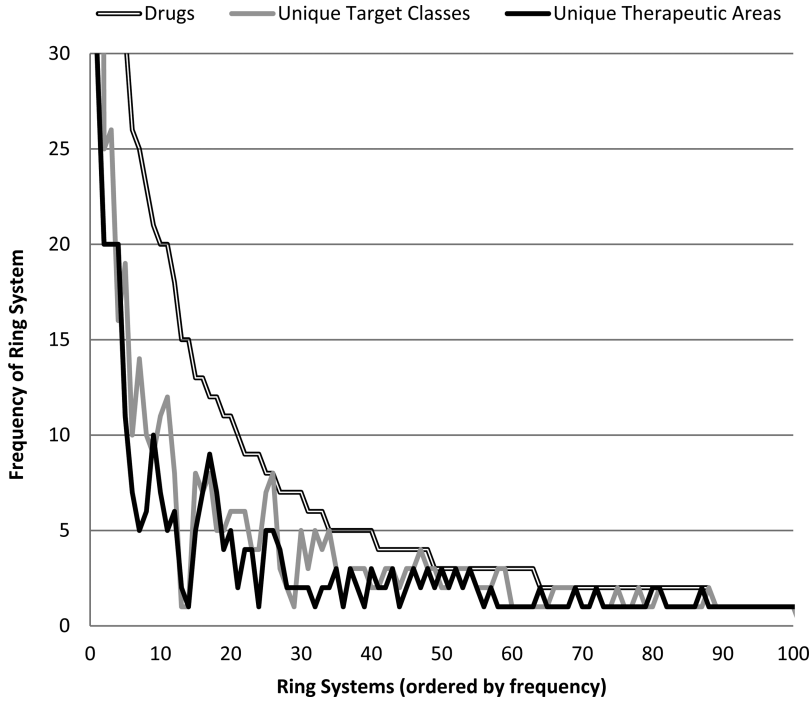


Figure 16. Application of ring systems across different therapeutic areas and target classes.

found in different therapeutic areas, e.g., kinases for inflammation and oncology, explaining the overlap. This is important for medicinal chemists who should be cognizant that ring systems cross both therapeutic areas and target classes. Although the overall acceptable toxicological profile may be different in different therapeutic areas, e.g., oncology versus dermatology, this may still be reflected in various rings being associated with specific toxicological profiles.

SUMMARY AND CONCLUSIONS

In this work we have analyzed the rings, ring systems, and frameworks of drugs as described in the FDA Orange Book until the end of 2012. A range of observations were covered and these are summarized below:

- Current drug space comprises only 351 ring systems and 1197 frameworks.
- There are over 204 ring systems and 901 frameworks that have only been used once in a drug.
- On average six new ring systems enter drug space per year.
- Each year, on average 28% of new drugs contain one new ring system, and therefore, 72% of new drugs will comprise only those ring systems found in previously marketed drugs.
- Less than 1% of drugs contain more than one new ring system.
- 67% of the ring systems that have been repeated in more recent years (after 1983) were first approved in drugs prior to 1983.
- From the top 100 most frequently used rings systems in all drugs, 83% are from drugs released prior to 1983.
- 42% of ring systems in drugs have been repeated, and from those 65% have been repeated within 10 years.
- 95% of drugs have five rings or less, and 99% have four ring systems or less.
- The overall count of O, N, and heteroatoms in a ring system is usually not greater than 3, 4, and 5 respectively.
- 40% of drugs do not contain any sp^3 carbons in a ring system.
- Ring systems can cross therapeutic areas and target classes. If a ring system is reused, 62% are for a different therapeutic area and 71% for a different target class.

These observations can be used as guidelines for atom counts, ring size, and frequencies for ring systems and frameworks. Moreover, on the basis of these observations, we can consider the importance of chemical novelty in drugs, although we note that this analysis focuses on chemical novelty as measured by the rings in drugs and compares this chemical innovation to successfully launched drugs. However, there are many other factors determining the success of drugs, many of which are nonscience related. For example, the peak in 1996 and 1997 is widely accepted to be in some part due to the PDUFA and the impact on review times and approval. The purpose of this legislation was to reduce the time and cost of drug development by facilitating a rapid review process which was funded by fees collected from New Drug Application (NDA) sponsors. Berndt et al.[38] suggested that without the PDUFA the peak in 1996 and 1997 would have been reduced, although whether the significant increase in approvals was solely attributed to this organizational change rather than scientific innovation is still questionable. Furthermore, this peak if taken in isolation could be used to suggest that productivity has dropped since 1996, although over the period from 1983 to 2012 the overall increase in new small molecule NMEs per year has remained relatively constant (22 per year) and the new chemical ring systems and frameworks per year has also remained constant.

Although this gives an indication of chemical novelty of drugs on the market, what we cannot determine from this analysis is whether the same distribution of ring systems and frameworks is observed in molecules currently in clinical trials. Clearly this gives an indication for successful candidates, but we cannot ascertain whether novel chemistries are more prevalent in clinical trials compared with marketed drugs or whether the ratio is consistent

dx.doi.org/10.1021/jm4017625 | J. Med. Chem. 2014, 57, 5845−5859

across clinical trials. We hope to address this question in future research.

This analysis does indicate the very small fraction of chemical space[44] and in particular ring system space covered by all currently available drugs. Moreover, only 28% of all new drugs contain one novel ring system suggesting that chemical novelty is not as important as bringing together the correct ring systems from drug space. If a novel ring system is brought to the market, then based on this analysis it is very unusual for the other ring systems in the drug to also be novel. Furthermore, it seems that medicinal chemistry relies heavily on a subset of ring systems that has not changed since at least 1982, which will in part be due to synthetic expediency.[45] However, on the basis of this observation, one could take the view that we do not introduce sufficiently new chemotypes into drug space which may be important for new drugs that target protein−protein interactions or allosteric modulation which are considered to be difficult, but important target types. An alternative perspective is that this small set of ring systems and frameworks are a pragmatic toolbox that are sufficient to take a drug to market.

We believe this analysis is useful in the design of hit finding libraries, in benchmarking libraries to assess the number of combinations of drug ring systems, and in the optimization of molecules. Since pre-existing ring systems account for 70% of drugs on the market, significant novelty can be meaningfully introduced from new combinations of known ring systems. However, novelty is often assessed by the scaffold, which can potentially ignore the need for novel configurations of current ring systems. Optimization of either an initial hit molecule or more elaborated lead can make use of the fact that if a scaffold already has a novel ring system, then evaluation of the known drug ring systems presented in this work may be advantageous before consideration of the introduction of additional novel ring systems. Moreover, when a small molecule is optimized, comprehensive coverage of the drug ring systems would provide valuable and extensive data from the perspectives of both intellectual property and biological activity.

■ ASSOCIATED CONTENT

Supporting Information

A3 pdf poster of the complete ring systems in drugs, sorted by frequency and then molecular weight. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: rich.taylor@ucb.com. Phone: +44(0)1753 807425.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare the following competing financial interest(s): Alastair D. G. Lawson is an employee of and holds shares and options in UCB. Richard D. Taylor is an employee of UCB.

Biographies

Richard D. Taylor obtained his M.Chem. in 1997 and Ph.D. (with Prof. Jonathan W. Essex on flexible docking of small molecules) in 2001 from the University of Southampton, U.K., supported by AstraZeneca. In 2001 he joined Astex Pharmaceuticals, where he was involved in virtual screening, docking software development, fragment library design, and fragment based drug discovery, working on successful oncology projects that have entered clinical trials. He moved to UCB in 2005 and is now a Principal Scientist in the Computer-Aided Drug Design group, involved in a range of drug discovery programs for inflammation and CNS disease areas. This has included leading a protein−protein interaction project for inflammation, and he currently leads the fragment technology project at UCB.

Malcolm MacCoss obtained his B.Sc. in Chemistry (1968) and his Ph.D. (1971, with Professors A. S. Jones and R. T. Walker) from the University of Birmingham, U.K. He then completed postdoctoral work at the University of Alberta, Canada, with Professor M. J. Robins. From 1976 to 1982, he worked at Argonne National Laboratory, U.S., on structural studies of nucleic acid components by NMR methods and novel prodrug approaches for nucleoside anticancer drugs. In 1982 he joined Merck Research Laboratories in Rahway, NJ, leaving in 2008 as Vice President for Basic Chemistry. He joined Schering-Plough as Group Vice President for Chemical Research and left in 2010 to found Bohicket Pharma Consulting LLC. He was appointed Visiting Professor of Chemistry for Medicine at the University of Oxford, U.K., in 2013.

Alastair D. G. Lawson was awarded a First in Biology at the University of Southampton in 1980. In 1983 he completed his Ph.D. at the Tenovus Research Laboratory at Southampton General Hospital, where he worked with Professor George Stevenson on anti-idiotype-based therapy for B cell leukemia. He joined Celltech as a research scientist in 1983 and has been closely involved with the discovery of many therapeutic antibodies, including Mylotarg, Cimzia, CMC-544, romosozumab, and olokizumab. Alastair led the development of UCB's proprietary antibody variable region discovery platform. As Research Fellow and Vice President of Structural Biology, he currently heads UCB's A2HiT initiative, in which information derived from antibodies is being applied to the discovery of new generations of small molecules addressing protein−protein interactions.

■ ACKNOWLEDGMENTS

We thank our colleagues at UCB for many helpful comments and discussions, including Luis Castro, Gillian Burgess, Johnny Zhu, Terry Baker, Chris Williams, and members of the UCB Computer-Aided Drug Design group including Florence Lebon, Mike King, Will Pitt, and Boris Kroeplien. We also thank Manoj C. Desai for providing the data from the *Annual Reports in Medicinal Chemistry*.

■ ABBREVIATIONS USED

CAS, Chemical Abstracts Service; PDUFA, Prescription Drug User Fee Act; sd, standard deviation

■ REFERENCES

(1) 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 Delivery Rev.* **2001**, *46*, 3−26.

(2) Sadowski, J.; Kubinyi, H. A Scoring Scheme for Discriminating between Drugs and Nondrugs. *J. Med. Chem.* **1998**, *41*, 3325−3329.

(3) Leeson, P. D.; Springthorpe, B. The Influence of Drug-like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discovery* **2007**, *6*, 881−890.

(4) Raju, T. N. The Nobel Chronicles. *Lancet* **2000**, *355*, 1022.

(5) Kola, I.; Landis, J. Can the Pharmaceutical Industry Reduce Attrition Rates? *Nat. Rev. Drug Discovery* **2004**, *3*, 711−715.

(6) Ajay; Walters, W. P.; Murcko, M. A. Can We Learn To Distinguish between "Drug-like" and "Nondrug-like" Molecules? *J. Med. Chem.* **1998**, *41*, 3314−3324.

(7) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615−2623.
(8) Leeson, P. D.; Davis, A. M. Time-Related Differences in the Physical Property Profiles of Oral Drugs. *J. Med. Chem.* **2004**, *47*, 6338−6348.
(9) Yusof, I.; Segall, M. D. Considering the Impact Drug-like Properties Have on the Chance of Success. *Drug Discovery Today* **2013**, *18*, 659−666.
(10) Gillet, V. J.; Willett, P.; Bradshaw, J. Identification of Biological Activity Profiles Using Substructural Analysis and Genetic Algorithms. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 165−179.
(11) Oprea, T. I. Current Trends in Lead Discovery: Are We Looking for the Appropriate Properties? *J. Comput.-Aided Mol. Des.* **2002**, *16*, 325−334.
(12) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A "Rule of Three" for Fragment-Based Lead Discovery? *Drug Discovery Today* **2003**, *8*, 876−877.
(13) Siegel, M. G.; Vieth, M. Drugs in Other Drugs: A New Look at Drugs as Fragments. *Drug Discovery Today* **2007**, *12*, 71−79.
(14) Dalvie, D.; Sajiv, N.; Kang, P.; Loi, C.-M. Influence of Aromatic Rings on ADME Properties of Drugs. In *Metabolism, Pharmacokinetics and Toxicity of Functional Groups*; Royal Society of Chemistry: Cambridge, U.K., 2010; pp 275−327.
(15) Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L. How To Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge. *Nat. Rev. Drug Discovery* **2010**, *9*, 203−214.
(16) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887−2893.
(17) Lewell, X. Q.; Jones, A. C.; Bruce, C. L.; Harper, G.; Jones, M. M.; McLay, I. M.; Bradshaw, J. Drug Rings Database with Web Interface. A Tool for Identifying Alternative Chemical Rings in Lead Discovery Programs. *J. Med. Chem.* **2003**, *46*, 3257−3274.
(18) Ertl, P.; Jelfs, S.; Mühlbacher, J.; Schuffenhauer, A.; Selzer, P. Quest for the Rings. In Silico Exploration of Ring Universe To Identify Novel Bioactive Heteroaromatic Scaffolds. *J. Med. Chem.* **2006**, *49*, 4568−4573.
(19) Wang, J.; Hou, T. Drug and Drug Candidate Building Block Analysis. *J. Chem. Inf. Model.* **2010**, *50*, 55−67.
(20) Lee, M. L.; Schneider, G. Scaffold Architecture and Pharmacophoric Properties of Natural Products and Trade Drugs: Application in the Design of Natural Product-Based Combinatorial Libraries. *J. Comb. Chem.* **2001**, *3*, 284−289.
(21) Walters, W. P.; Green, J.; Weiss, J. R.; Murcko, M. A. What Do Medicinal Chemists Actually Make? A 50-Year Retrospective. *J. Med. Chem.* **2011**, *54*, 6405−6416.
(22) Miyaura, N.; Yamada, K.; Suzuki, A. A New Stereospecific Cross-Coupling by the Palladium-Catalyzed Reaction of 1-Alkenylboranes with 1-Alkenyl or 1-Alkynyl Halides. *Tetrahedron Lett.* **1979**, *20*, 3437−3440.
(23) Ertl, P. Database of Bioactive Ring Systems with Calculated Properties and Its Use in Bioisosteric Design and Scaffold Hopping. *Bioorg. Med. Chem.* **2012**, *20*, 5436−5442.
(24) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40*, D1100−D1107.
(25) Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.; Djoumbou, Y.; Eisner, R.; Guo, A. C.; Wishart, D. S. DrugBank 3.0: A Comprehensive Resource for "Omics" Research on Drugs. *Nucleic Acids Res.* **2011**, *39*, D1035−D1041.
(26) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool To Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757−1768.
(27) Pitt, W. R.; Parry, D. M.; Perry, B. G.; Groom, C. R. Heteroaromatic Rings of the Future. *J. Med. Chem.* **2009**, *52*, 2952−2963.
(28) Lipkus, A. H.; Yuan, Q.; Lucas, K. A.; Funk, S. A.; Bartelt, W. F., 3rd; Schenck, R. J.; Trippe, A. J. CAS Registry. Structural Diversity of Organic Chemistry. A Scaffold Analysis of the CAS Registry. *J. Org. Chem.* **2008**, *73*, 4443−4451.
(29) Gibson, S.; McGuire, R.; Rees, D. C. Principal Components Describing Biological Activities and Molecular Diversity of Heterocyclic Aromatic Ring Fragments. *J. Med. Chem.* **1996**, *39*, 4065−4072.
(30) Ritchie, T. J.; Macdonald, S. J. F. The Impact of Aromatic Ring Count on Compound Developability—Are Too Many Aromatic Rings a Liability in Drug Design? *Drug Discovery Today* **2009**, *14*, 1011−1020.
(31) Young, R. J.; Green, D. V. S.; Luscombe, C. N.; Hill, A. P. Getting Physical in Drug Discovery II: The Impact of Chromatographic Hydrophobicity Measurements and Aromaticity. *Drug Discovery Today* **2011**, *16*, 822−830.
(32) FDA Home Page. http://www.fda.gov (accessed Aug 1, 2013).
(33) Wikipedia Home Page. http://www.wikipedia.org (accessed Aug 1, 2013).
(34) Mullard, A. 2012 FDA Drug Approvals. *Nat. Rev. Drug Discovery* **2013**, *12*, 87−90.
(35) Cumulative NCE Introduction Index, 1983−2011 (by Indication). In *Annual Reports in Medicinal Chemistry*; Desai, M. C., Ed.; Academic Press: San Diego, CA, 2012; Vol. *47*, pp 629−652.
(36) Accelrys Software Inc. http://www.accelrys.com.
(37) Munos, B. Lessons from 60 Years of Pharmaceutical Innovation. *Nat. Rev. Drug Discovery* **2009**, *8*, 959−968.
(38) Berndt, E. R.; Gottschalk, A. H. B.; Philipson, T. J.; Strobeck, M. W. Industry Funding of the FDA: Effects of PDUFA on Approval Times and Withdrawal Rates. *Nat. Rev. Drug Discovery* **2005**, *4*, 545−554.
(39) Kendall, M. G. A New Measure of Rank Correlation. *Biometrika* **1938**, *30*, 81−93.
(40) Leeson, P. D.; St-Gallay, S. A.; Wenlock, M. C. Impact of Ion Class and Time on Oral Drug Molecular Properties. *MedChemComm* **2011**, *2*, 91−105.
(41) Lovering, F.; Bikker, J.; Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **2009**, *52*, 6752−6756.
(42) Kenny, P. W.; Montanari, C. A. Inflation of Correlation in the Pursuit of Drug-likeness. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 1−13.
(43) Hung, A. W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J. A.; Clemons, P. A.; Young, D. W. Route to Three-Dimensional Fragments Using Diversity-Oriented Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6799−6804.
(44) Dobson, C. M. Chemical Space and Biology. *Nature* **2004**, *432*, 824−828.
(45) 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**, *54*, 3451−3479.