Journal of Medicinal Chemistry

PERSPECTIVE

pubs.acs.org/jmc

What Do Medicinal Chemists Actually Make? A 50-Year Retrospective

W. Patrick Walters,* Jeremy Green, Jonathan R. Weiss, and Mark A. Murcko

Vertex Pharmaceuticals Inc., 130 Waverly Street, Cambridge, Massachusetts 02139, United States

Supporting Information

INTRODUCTION

Despite the dramatically increasing costs of pharmaceutical R&D,[1−3] productivity, as measured by the number of new drugs approved each year, has languished during the past decade. Many different explanations have been proposed to account for this decline, such as increasing regulatory scrutiny,[4,5] the inherent inefficiency of very large organizations,[6,7] the pursuit of highly speculative targets identified from genomics,[3] and more objective decision-making about advancing marginal compounds into late-stage clinical trials.[8]

Another factor often cited to help explain low pharmaceutical productivity is that, paradoxically, many compounds synthesized by medicinal chemists are not "druglike", at least as measured by various criteria such as the "rule of 5"[9] or REOS.[10] This finding can be rationalized in many different ways, as described below. To better understand which factors exert the greatest influence on the practice of medicinal chemistry, we speculated that it might be instructive to examine changes over time in the properties of molecules that medicinal chemists make. Toward this end, we have analyzed molecules reported in *Journal of Medicinal Chemistry* between Volumes 1 and 52 (1959−2009). A total of 415 284 molecules are included in this assessment.

Since 1959, *Journal of Medicinal Chemistry* (JMC) has been a primary source for reports from academic and industrial medicinal chemistry groups. It is the most cited journal in the field of medicinal chemistry with more than 42 000 citations recorded in 2009.[11] While a number of groups have published analyses of marketed drugs or compounds from recent pharmaceutical patents, we are unaware of reports that provide a similar analysis of the vast medicinal chemistry literature. Analyses of drugs do not necessarily represent the much larger fraction of medicinal chemistry compounds that do make it to market. Estimates of the ratio of medicinal chemistry compounds synthesized to approved drugs vary widely. Medicinal chemistry groups in large pharmaceutical companies synthesize in excess of 100 000 compounds annually; yet these companies typically receive approval for only a few new drugs each year. Other organizations may claim higher productivity, but we can comfortably assume that well in excess of 10 000 medicinal chemistry compounds are synthesized for each approved drug. While it is obvious that not every compound synthesized in a drug discovery program becomes part of a publication, we can assume that publications in JMC provide a significantly broader overview of the field than the small number of drugs that are either approved or enter clinical trials every year. As a means of approximating the fraction of medicinal chemistry output represented by JMC, we calculated the similarity of JMC compounds to the output from medicinal chemistry programs at Vertex Pharmaceuticals between 1990 and 2010. Approximately 5% of Vertex compounds had at least a 0.7 Tanimoto similarity (Daylight fingerprints) to at least one JMC compound. This fraction may be greater for larger organizations. In spite of the fact that JMC coverage is not comprehensive, we feel that it provides an adequate representation of the overall trends in medicinal chemistry.

Over the past 10 years, a number of groups have carried out analyses of the properties of marketed drugs and compounds under clinical investigation. A thorough summary of this work can be found in a 2004 review by Lajiness.[12] A 2004 paper by Vieth showed that there were minimal changes in the median molecular weight and CLogP of drugs launched between 1982 and 2002.[13] However, different conclusions can be reached when drugs are grouped chronologically and different time periods are considered. In 2004, Leeson and Davis[14] published an analysis that examined the difference between oral drugs launched prior to 1983 (864 drugs) and those launched between 1984 and 2004 (329 drugs). They found that the mean molecular weight, numbers of oxygen and nitrogen atoms, the number of rings, and the number of rotatable bonds had increased. However, there appeared to be no statistically significant difference in CLogP, % polar surface area, and number of hydrogen bond donors between the pre- and post-1984 drugs.

A pair of papers from Leeson and co-workers compared calculated properties of compounds taken from recent patents published by major pharmaceutical companies with those of a set of marketed oral drugs. In a 2007 paper,[15] the authors found that molecular weight and number of hydrogen bonding groups in drugs had increased significantly over time, while increases in CLogP for drugs had been more modest. The authors also found that compounds in recent pharmaceutical patents were larger (median MW of 451 vs 432) and more lipophilic (median CLogP of 4.1 vs 3.1) than drugs approved since 1990. In a 2010 publication,[16] the authors examined the relationship between calculated properties and ion class over time. They found that, with the exception of acidic drugs approved after 1990, median values for CLogP, polar surface area, and number of hydrogen bond donors for drugs had been relatively constant since the 1960s. While the number of hydrogen bond acceptors in drugs increased, there were no time-dependent trends for shape or chirality. The authors again found that compounds in current pharmaceutical patents tend to be larger, more lipophilic, and less three-dimensional than oral drugs. An examination of compounds appearing in pharmaceutical patents during 2000−2009 showed no trends over time for molecular weight, CLogP, and shape, and the authors propose that inflation of these properties may be leveling.

Received: April 25, 2011
Published: July 14, 2011

dx.doi.org/10.1021/jm200504p | *J. Med. Chem.* 2011, 54, 6405–6416

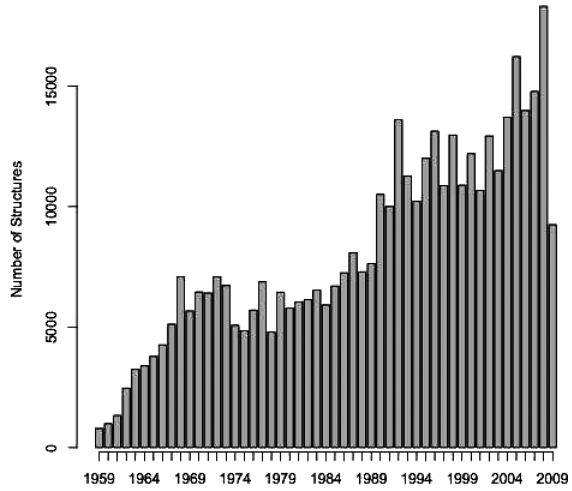

Figure 1. Number of structures per year from *Journal of Medicinal Chemistry* in the combined database used for this study. Records were only available for the first half of 2009.

MATERIALS AND METHODS

Data used in this study was obtained by combining data sets from two databases, the MedChem database (GVK Biosciences, Hyderabad, India) and the publicly available ChEMBL[17−19] database from EMBL. Chemical structures and year of publication for 408 330 compounds from 18 629 articles published in *Journal of Medicinal Chemistry* between 1959 and 2007 were extracted from the MedChem database as an SD file. Chemical structures for 335 640 compounds from 14 447 JMC papers published between 1980 and 2009 were extracted as an SD file from version 2 of the ChEMBL database. SMILES strings were generated for canonical tautomers of each molecule using the QuacPac toolkit (OpenEye Scientific Software, Santa Fe, NM) and used to create a set of 422 678 records corresponding to unique combinations of SMILES, volume, and starting page. 7645 chemical structures with a molecular weight greater than 1000 were excluded, leaving a total of 415 284 structures from 19 299 papers that were used for the present analysis. A calculation of the overlap between the JMC set of 415K structures and 2009 edition of the MDL Drug Data Report (MDDR) (Symyx, San Leandro, CA) indicated that compounds used in this analysis contained 512 launched drugs and 258 compounds that had been in clinical trials. The number of structures in the database on a per year basis is presented in Figure 1. Note that the records were only available for the first half of 2009 in the version of the ChEMBL database used for this analysis. Summary statistics for the number of structures per article are shown in Table 1.

For comparison, a set of 1219 marketed drugs with molecular weight less than 1000 was extracted from the 2009 edition of the MDDR database. All entries with the "PHASE" field set to "Launched" were selected. Year of approval for the drugs was taken from the earliest entry in the "PHASE_YEAR" field. The same properties calculated for the JMC set were calculated for each of the drugs. These properties are depicted as dotted lines in Figures 2 and 3, with the 95% confidence interval about the mean plotted as a gray region.

Calculations of molecular weight, number of hydrogen bonding groups, and rotatable bonds were carried out using internally developed Python scripts that employ the OEChem toolkit (OpenEye Scientific Software, Santa Fe, NM). The SMARTS patterns in Table 2 were used to identify these groups.

Table 1. Summary Statistics for the Number of Compounds per Article in the Combined Database of Compounds Appearing in JMC between 1959 and 2009

	no. compds per article
minimum	1
first quartile	7
median	14
mean	21.5
third quartile	28
maximum	1357

Calculation of polar surface area was performed using an internally developed C++ program that uses the method originally published by Ertl.[20] This program also uses the OEChem toolkit for all molecule-handling functions. The log *P* calculations (CLogP) were performed using an internally developed and validated implementation of the XLogP algorithm.[21,22] Molecular complexity was assessed using the MDL 166 bit fingerprints[23] as implemented in the OEGraphSim toolkit from OpenEye Scientific Software.

All statistical analyses and graphics were generated with R, version 2.11.1. The plots in Figures 2 and 3 show the mean and median values for compounds appearing in papers published during a specific 5-year interval as solid lines. Mean values are depicted as colored boxes with a blue box representing a statistically significant difference (p for difference in means is ≤ 0.05) with the preceding interval and a red box representing a lack of statistical significance (p for difference in means is >0.05). Calculations of statistical significance were performed using the R function "pairwise.t.test". This function performs multiple pairwise comparisons and uses Holm's method[24] to correct for multiple testing.

MOLECULAR WEIGHT

Molecular weight has been correlated with a number of key parameters in drug discovery. It has been reported that over the past few decades, the molecular weights of marketed drugs as well as drug candidates have been increasing.[25] Lipinski[26] has shown that there was a marked increase in molecular weight for clinical candidates put forward by both Merck and Pfizer between 1962 and 1997. Leeson and Davis[14] showed that molecular weights of marketed oral drugs approved between 1983 and 2002 had a molecular weight 14% greater than those approved before 1983. In a 2008 study, Gleeson[27] used principal components analysis to identify an optimal set of descriptors to use in the generation of predictive models for 12 ADME assays run routinely at GSK. Molecular weight was the most significant of the descriptors tested and explained 92% of the variation in the first principal component. These reports, coupled with a high number of failures in early clinical trials, have brought about a renewed focus on physical properties with a particular emphasis on lower molecular weight compounds. The revival of this interest seems to have been driven, in part, by Lipinski's publication of the "rule of 5",[9] which helped to define property ranges that have become a common part of many drug discovery efforts. There has also been a renewed interest in leadlikeness which has driven the recent growth of fragment-based design and strategies such as ligand efficiency[28−30] and the "rule of 3".[31]

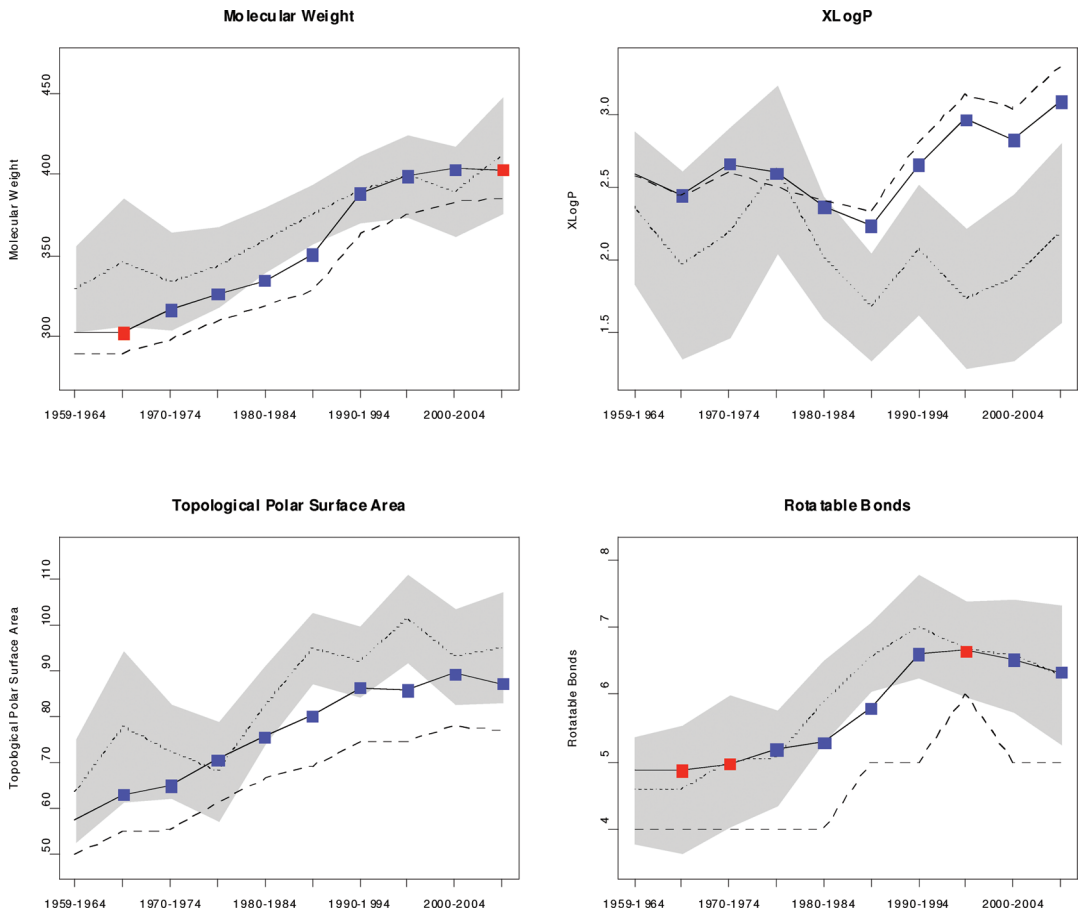

Figure 2. Mean values for molecular weight, CLogP, polar surface area, and number of rotatable bonds for molecules published in *Journal of Medicinal Chemistry* in 1959−2009 are shown as solid lines with squares. Blue squares indicate a statistically significant difference between the mean for a particular 5-year interval and the preceding interval. Red squares indicate that the difference was not statistically significant. Dashed lines indicate the median value for each property during the designated interval. The gray regions represent the 95% confidence interval for drugs launched during the same time period, with the dotted line representing the mean value for drugs during each 5-year period.

The relationship between molecular weight and selectivity is still in question. A recent study from Pfizer[32] showed that higher molecular weight molecules tend to be more selective, while a similar study from Novartis[33] reached the opposite conclusion that larger molecules tend to be more promiscuous. A third study,[34] published this year, indicates that for kinase inhibitors synthesized at BMS, molecular weight is not a determining factor for selectivity. Differences in targets and pharmacology may make it difficult to generalize something as complex as selectivity.

Figure 2 shows the molecular weight trends for compounds appearing in JMC papers between 1959 and 2009, divided into 5-year intervals. Trends for JMC compounds parallel those seen for drugs. For JMC compounds, there is a statistically significant increase in both the mean and median over time. During the 1960s, the median molecular weight of compounds in JMC was close to 300 (303 for both 1959−1964 and 1965−1969). In the 1970s, the median molecular weight increased by 2−3% over each 5-year period, rising to 317 for 1970−1974 and 326 for 1975−1979. The increase in molecular weight continued at a similar rate throughout the 1980s. The most significant increase in molecular weight occurred in the period between 1990 and 1994 with the median molecular weight increasing 10% over the previous 5-year period. In the time between 1995 and 2009, increases in molecular weight were more modest, with increases of 0−2% over intervening 5-year periods with a median value of 388 for the most recent period, 2005−2009. In aggregate, the molecular weight of the average JMC compound has increased by 25% during the entire 50-year period.

CLOGP

Lipophilicity is another key factor in drug design.[35] The most common surrogate for lipophilicity is the octanol−water partition coefficient or log *P*. Computational predictions of log *P* have been studied for more than 40 years[36] and have become a routine part of medicinal chemistry practice. Calculated log *P* (CLogP) values have been correlated with a wide array of key drug discovery parameters including solubility,[37−39] permeability,[40,41] hERG binding,[35,42,43] and CNS penetration.[44,45] It has also been noted that lipophilic compounds tend to be more promiscuous and have a higher incidence of nonspecific toxicity.[46] Changes in calculated log *P* for drugs over time have received some attention in the recent literature.[14,25] Lipinski reported that the CLogP for compounds synthesized in drug discovery programs at Pfizer in 1994 was 14% higher than for those synthesized in 1986.[26]

Figure 2 shows CLogP trends for drugs and compounds published in JMC over the same time intervals as the previous analysis. It is interesting to note that since 1985, the mean CLogP of compounds published in JMC has consistently been higher than

dx.doi.org/10.1021/jm200504p |*J. Med. Chem.* 2011, 54, 6405–6416

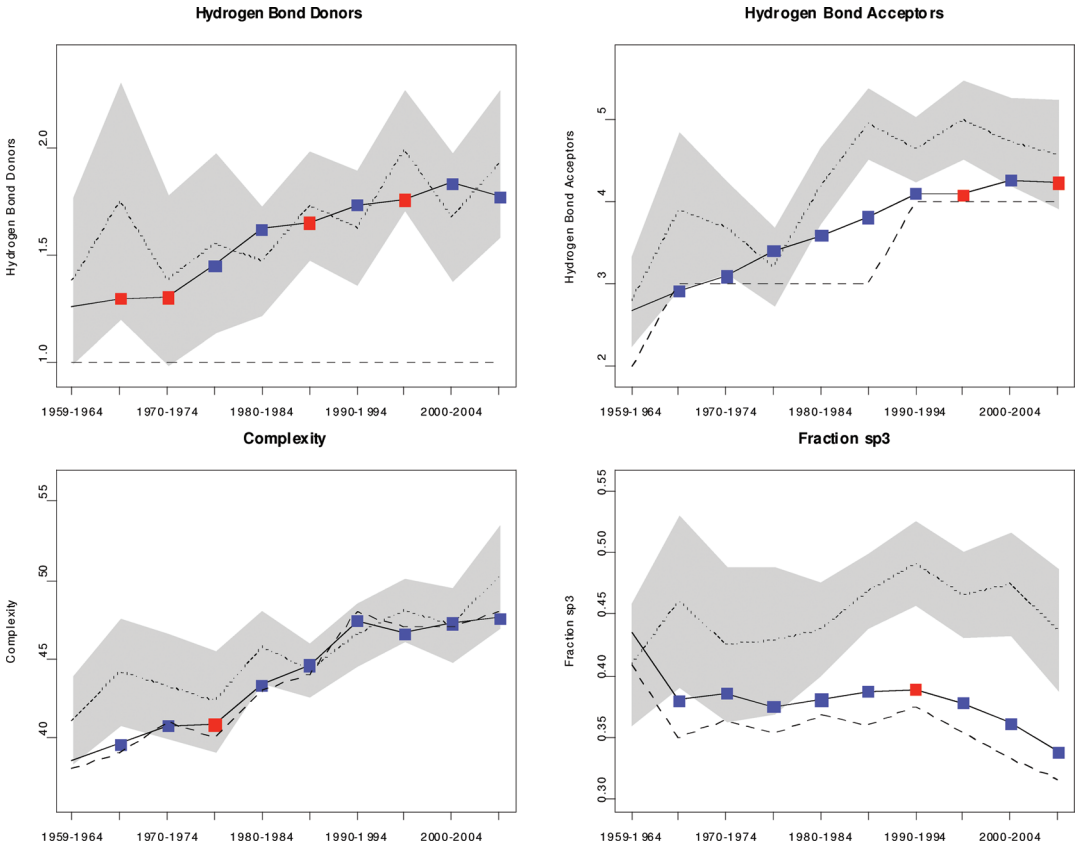

Figure 3. Mean values for number of hydrogen bond donors and acceptors, molecular complexity, and fraction of sp^3 atoms for molecules published in *Journal of Medicinal Chemistry* in 1959−2009. Blue squares indicate a statistically significant difference between the mean for a particular 5-year interval and the preceding interval. Red squares indicate that the difference was not statistically significant. Dashed lines indicate the median value for each property during the designated interval. The gray regions represent the 95% confidence interval for drugs launched during the same time period, with the dotted line representing the mean value for drugs during each 5-year period.

Table 2. SMARTS Patterns

	SMARTS pattern
hydrogen bond donor	[#7H,#7H2,#7H3,OH]
hydrogen bond acceptor	[#8H0,#8H1,#7H0&!$([#7X3,#7X4])
rotatable bond	[!$([NH]!@C(=O))&!D1&!$(*#*)]&!@[!$([NH]!@C(=O))&!D1&!$(*#*)]

that for current marketed drugs. Prior to 1985 there is no overall trend, with CLogP values fluctuating up or down by 2−4% over 5-year intervals. Starting in 1990, there was a dramatic increase in CLogP. Between 1990 and 1994 there was a 17% increase in the median CLogP value. This was followed by an additional 10% increase between 1995 and 1999. There was a smaller change between 2000 and 2004 (decrease of 2%), but CLogP values do appear to be on the increase again in the second half of the decade. However, because of the smaller sample, it is difficult to make a definitive assessment of the period between 2005 and 2009.

POLAR SURFACE AREA

Over the past 10 years, polar surface area (PSA) has received increased attention in the medicinal chemistry literature as well as in the pharmaceutical industry. It has been noted that molecules possessing a large PSA may encounter difficulty in transiting biological membranes.[47] This inability to cross membranes may result in poor absorption or a lack of blood−brain barrier (BBB) penetration. Some of the first work correlating PSA and oral absorption was published by Palm[48] who found that drugs that were highly absorbed (>90%) had a polar surface area less than 60 $Å^2$ while drugs that were poorly absorbed (<10%) had a polar surface area greater than 140 $Å^2$. Clark and others extended this work by employing larger data sets and alternative methods for computing polar surface area. Subsequent work by Egan[41] defined an ellipsoidal region in a property space defined by log *P* and PSA where compounds had a higher probability of being orally absorbed. This region or "egg", defined through an analysis of drugs in the Physician's Desk Reference, can then serve as a target for compound design. The relationship between PSA and BBB penetration was first published by van de Waterbeemd in 1992[49] and further refined by Clark and others in subsequent papers. Hitchcock and Pennington[50] suggest a cutoff of 90 $Å^2$ for compounds targeting the CNS and point out that the average PSA for the top 25 CNS drugs is 47 $Å^2$.

Figure 2 shows a consistent increase in both the mean and median polar surface area over time for drugs and compounds

dx.doi.org/10.1021/jm200504p |*J. Med. Chem.* 2011, 54, 6405–6416

published in JMC. PSA steadily rose over each of the 5-year intervals in the period between 1959 and 1995, with increases in the median value between 2% and 10% in each interval. With a couple of exceptions, PSA values for JMC compounds have been within the ranges observed for marketed drugs.

ROTATABLE BONDS

Molecular flexibility is another property that is frequently optimized over the course of drug discovery programs. Medicinal chemists often rigidify a molecule to reduce conformational entropy and increase binding affinity or impart selectivity. The number of rotatable (typically nonring single) bonds is often used as a surrogate for molecular flexibility. A 2002 paper by Veber[51] correlated rat oral bioavailability with number of rotatable bonds, polar surface area, and number of hydrogen bonding groups. Subsequent work by Lu[52] demonstrated that these relationships were more complex and results tended to vary based on calculation methods and the therapeutic class of the molecules being analyzed. A 2009 paper by Hou[53] cast further doubt on these results by analyzing a set of 706 compounds and finding no relationship between human oral bioavailability and rotatable bonds or any other calculated parameter.

Figure 2 shows an increase in the average number of rotatable bonds for compounds published in JMC between 1959 and 2009. These changes parallel the trends observed for marketed drugs. The mean value increased from 1959 to 1990 but appears to have stabilized over the past 20 years. The median number of rotatable bonds in JMC compounds was constant at 4 from 1965 through 1984. In the interval between 1985 and 1989, the number increased from 4 to 5, with a further increase to 6 in the second half of the 1990s.

HYDROGEN BONDING GROUPS

Hydrogen bonds are a key element of almost every drug discovery program. A properly placed hydrogen bond can impart both potency and selectivity to a compound. However, like many of the parameters discussed here, hydrogen bonds can be a double-edged sword. An excessive number of hydrogen bonding groups can inhibit a compound's ability to transit a biological membrane and significantly reduce bioavailability. In the 1997 "rule of 5" paper,[9] Lipinski noted that the majority of marketed drugs possessed fewer than 5 hydrogen bond donors and fewer than 10 hydrogen bond acceptors. These observations, which have been supported in a number of subsequent studies,[27,54−56] have become guidelines for many drug discovery programs.

Figure 3 shows trends for hydrogen bond donors and acceptors in molecules published in JMC between 1959 and 2009. The mean number of hydrogen bonding groups increased through the 1960s, 1970s, and 1980s but appears to have stabilized over the past 20 years. Trends for hydrogen bond donors paralleled those observed for marketed drugs, while the mean number of acceptors dropped below that observed for drugs during the 1980s and 1990s. The median number of hydrogen bond donors has been constant at 1 throughout the entire 50-year period. The number of hydrogen bond acceptors remained constant at 3 throughout the 1960s, 1970s, and early 1980s with an increase to 4 in the 1980s.

MOLECULAR COMPLEXITY

Over the past 50 years, medicinal chemists have attacked increasingly challenging targets.[57−59] In addition, closely related targets such as kinases or ion channels present selectivity challenges. One school of thought holds that more challenging targets may require additional functionality. In evaluating changes in medicinal chemistry compounds over time, we wanted a metric to compare the relative complexity of molecules. We chose to employ an approach used in a publication by Schuffenhauer,[60] who quantified molecular complexity based on the number of bits turned on in a molecular fingerprint.

Table 3. Number of Rings, Linkers, and Side Chains Required To Describe 415 000 Compounds Published in JMC between 1959 and 2009

	70% coverage	80% coverage	90% coverage	total
rings	37	112	470	12172
linkers	53	127	518	8881
side chains	16	40	265	25330

Figure 3 provides a comparison of complexity for molecules reported in JMC between 1959 and 2009. As with many of the other parameters reported here, complexity increased steadily for both drugs and JMC compounds between 1959 and 1990. The median complexity for JMC compounds increased in the 1980s with an 8% increase for the period 1980−1985 and an additional 5% increase in the subsequent 5-year period. JMC complexity continued to increase by 10% between 1990 and 1995 but has leveled off in the intervening years. However, perhaps because of an increasing need for selectivity and safety, complexity of marketed drugs increased between 2004 and 2009.

"FLATNESS"

The aromatic character of druglike molecules has received a significant amount of recent attention. One measure of a molecule's aromatic character, and indirectly of its three-dimensional shape, is the fractional sp^3 character (Fsp3). This metric, originally defined by Yan and Gasteiger,[61] is defined as the ratio of the number of sp^3 carbon atoms to the total number of carbon atoms. A 2009 publication by Lovering and co-workers[62] showed that Fsp3 can be related to physical properties such as aqueous solubility and melting point. The authors demonstrate that molecules with a higher Fsp3 tend to have lower melting points and higher aqueous solubility. These properties would potentially lead to molecules that would be easier to formulate. The authors also showed that marketed drugs tend to have a higher Fsp3 (0.47) than discovery compounds (0.36). Recent work by Ritchie and McDonald[63] has shown a relationship between the number of aromatic rings and several properties such as solubility, CYP inhibition, plasma−protein binding, and hERG binding, all of which influence the challenges inherent in developing a compound.

Figure 3 shows the trends for Fsp3 for molecules published in JMC between 1959 and 2009. As can be seen in the figure, the median fraction of sp^3 carbons increased from 0.35 in the mid-1960s to 0.38 in the early 1990s. However, Fsp3 steadily decreased between 1995 and 2009, reaching a median value of 0.32 for the most recent period measured. It is interesting to note that since the mid-1970s, molecules published in JMC have consistently had less sp^3 character than marketed drugs. As mentioned by Lovering,[62] the increasing "flatness" of molecules may have come about because of new chemistries that facilitate sp^2-sp^2 couplings.

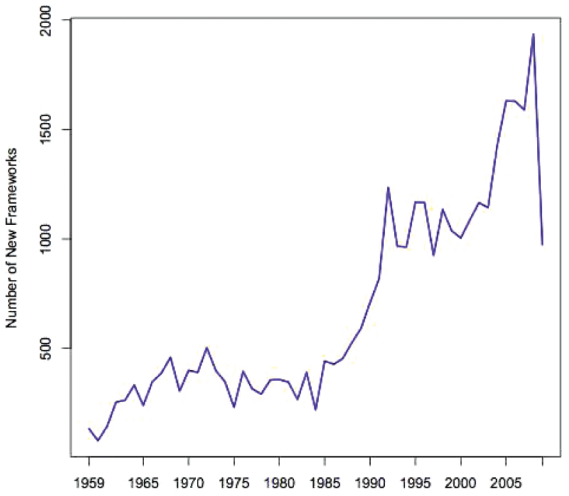

Figure 4. Number of new molecular frameworks published per year between 1959 and 2009. Records were only available for the first half of 2009.

MOLECULAR FRAMEWORKS

It is also interesting to note (Table 3) that 70% of the compounds in JMC are composed of a relatively small number of "building blocks": 37 ring types, 53 linkers, and 16 side chains. The data in Table 3 were generated by fragmenting molecules according to the scheme originally described by Bemis;[64] rings, linkers, and side chains were sorted by frequency. We then counted the number of rings, linkers, or side chains required to account for 70%, 80%, and 90% of the total. This finding is consistent with earlier studies of known drugs.[64,65] On the other hand, there is strong evidence (Figure 4) that the rate of introduction of new molecular frameworks (combinations of rings + linkers) has actually risen in recent years, with 1589 in 2007 and 1935 in 2008. In other words, the same building blocks are being assembled in novel ways. It is interesting to note that the number of new frameworks seems to parallel an increased number of acyclic−aromatic bonds in JMC molecules (vide infra). 25% of the new frameworks published in 2008 contain at least one acyclic−aromatic bond. It appears that the increase in the number of new frameworks may be driven, at least in part, by the adoption of palladium catalyzed couplings such as the Suzuki,[66] Heck,[67] and Negishi[68] reactions.

RESULTS AND DISCUSSION

Our analysis of more than 400 000 molecules reported in *Journal of Medicinal Chemistry* during the past 50 years clearly shows that the properties of the typical molecule have changed dramatically. Molecules are getting larger, more complex, and more lipophilic, and they are flatter and more aromatic, as measured by the fraction of sp3 carbons. In addition these molecules are more flexible, they have an increased polar surface area, and they contain more hydrogen bond donors and acceptors. Not surprisingly, many of these trends parallel those observed for marketed drugs. Table 4 and Figure 6 provide comparisons of properties calculated for molecules published in the first 5 years of JMC with compounds published in the most recent 5-year period. As can be seen in Table 4, mean values for all calculated properties increased by 16−37% except for the fraction sp^3 which decreased by 28%. Figure 6 shows distributions of the same sets as box plots.

Table 4. Mean Properties for Compounds Appearing in Papers during the First 5 Years of JMC Publication and in the Most Recent 5-Year Period[a]

	1959−1964	2005−2009	% difference
molecular wt	302.68	403.29	**24.9**
CLogP	2.60	3.10	**16.0**
TPSA	57.50	87.05	**33.9**
rotatable bonds	4.86	6.33	**23.2**
HB donors	1.25	1.77	**29.3**
HB acceptors	2.67	4.24	**37.0**
complexity	38.51	47.59	**19.1**
fraction sp^3	0.43	0.34	**−28.5**

[a] Values in bold indicate a statistically significant difference in the mean ($p < 0.05$).

The distributions for drugs launched during the same intervals are shown for reference. Notches in the boxplots provide a rough idea of the 95% confidence interval about the median. If the notches about two medians do not overlap, there is strong evidence that medians do not overlap.[69] Median values for all properties with the exception of fraction sp^3 increased over the past 50 years. Fraction sp^3 decreased significantly over the same period. Further examination of Figure 6 indicates that the distributions of CLogP and the fraction sp^3 for molecules published in the most recent 5-year period lie outside those observed for drugs launched during the same interval.

Most chemists would interpret these data to suggest that some of the properties of molecules are getting "worse" or what is sometimes called "less druglike". While the interpretation may be debatable, it is unquestionably true that the CLogP and fraction sp^3 of drugs and typical JMC molecules, over time, have increasingly diverged; that is, a higher percentage of the JMC molecules have properties that fall outside the ranges most commonly found in successful drugs. Of course, this does not necessarily mean that these are "bad" molecules; many successful drugs fall outside the normal ranges.

These trends have been evident for decades but may be abating. In Figures 2 and 3, most of the curves appear to have leveled over the past 15 years. In Table 5, we compare the mean values from 1959 to 1999 to those of 2000−2009. In every category except Fsp3, the more recent compounds have higher mean values. Consistent with other properties moving away from "druglikeness", Fsp3 decreased. However, when we further compare the most recent 5 years (2005−2009) to the preceding 5 years (2000−2004), we see many trends moving back in "druglike" directions. However, one has to be careful not to overinterpret short-term trends.

The fact that properties of marketed drugs have changed over time calls into question our notions of druglikeness. When Lipinski published the "rule of 5" paper[9] in 1997, the average molecular weight of launched drugs was 369. During the most recent 5-year period, the average molecular weight was 388, an increase of 6%. Should we make adjustments to bring the rule of 5 in line with current practice, or should we adopt a term other than druglike that more adequately describes the desirable properties of molecules?

To better understand how we might rethink the concept of "druglike," we can think of nine possible explanations for changes in the properties of molecules made by medicinal chemists over the past 50 years.

Table 5. Mean Properties for Compounds in JMC before 2000 and in the Subsequent 10-Year Period, as Well as Differences for the Two Most Recent 5-Year Periods[a]

	1959−1999	2000−2009	% difference	2000−2004	2005−2009	% difference
molecular wt	354.09	403.35	**13.91**	403.43	403.29	−0.04
CLogP	2.61	2.98	**13.84**	2.83	3.1	**9.26**
TPSA	76.76	88.08	**14.74**	89.29	87.05	**-2.51**
rotatable bonds	5.80	6.42	**10.56**	6.51	6.33	**-2.8**
HB donors	1.58	1.80	**14.26**	1.83	1.77	**-3.36**
HB acceptors	3.65	4.25	**16.31**	4.27	4.24	−0.65
complexity	43.91	47.45	**8.05**	47.28	47.59	**0.64**
fraction sp^3	0.39	0.35	**−9.37**	0.36	0.34	**−6.54**

[a] Values in bold indicate a statistically significant difference in the mean ($p < 0.05$).

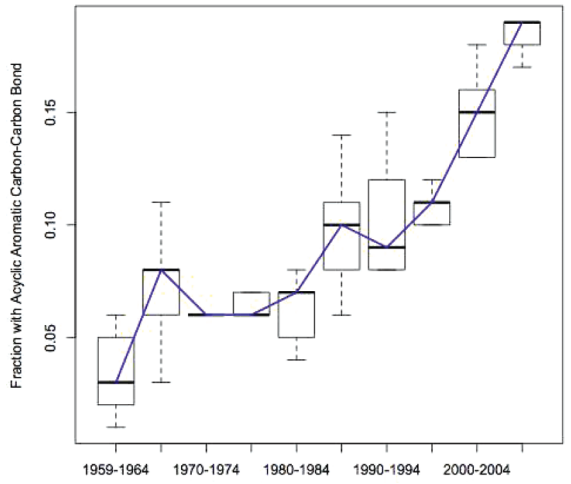

Figure 5. Influence of sp^2-sp^2 coupling chemistries on the molecules published in JMC between 1959 and 2009. Data are shown as the fraction of molecules published in each 5-year period containing at least one acyclic−aromatic carbon−carbon bond.

Advances in Molecular Biology. With increasing knowledge of receptor subtypes, and therefore greater concerns about selectivity, came an increase in target-based counterscreening.[70−72] Achieving selectivity in practice frequently requires larger or more complex molecules.

Advances in Synthetic and Analytical Techniques. It is far easier today than in the 1970s to make, purify, and structurally characterize complex molecules. Therefore, quite naturally, there is a tendency to take advantage of these improvements to make more complex molecules. The increased availability of reagents also makes it much easier to synthesize larger, more complex molecules. As an example, consider arylboronic acids, a common component in the widely used Suzuki coupling. Over the past 15 years, Vertex Pharmaceuticals has maintained an extensive database of available reagents. In 2000, this database contained approximately 300 commercially available arylboronic acids. A search in the current version of the same database yields more than 3000 such reagents.

Advances in High-Throughput Chemistries. As mentioned above, the introduction of new methods for sp^2-sp^2 couplings, and the adaptation of these methods to high-throughput synthesis, has had a broad impact on medicinal chemistry. Figure 5 shows the fraction of molecules containing at least one acyclic single bond connecting two aromatic rings in molecules published in JMC between 1959 and 2009. The marked increase in the fraction of molecules with acyclic bonds connecting aromatic rings over the past 15 years is consistent with the increase in "flatness" (decrease in fraction sp^3).

Better Formulation Technologies Can Precipitate a Sense of Complacency. Because formulations scientists are so often able to dramatically improve the pharmacokinetic properties of molecules,[73−75] there may be a tendency for medicinal chemists to drive forward with suboptimal compounds, expecting that physicochemical deficiencies will be overcome in development.

Target-Focused Drug Discovery. Increasing application of molecular understanding of disease has led to target-based research, which facilitates sequential optimization of properties. For example, discovery teams often optimize receptor binding, then cellular activity, then ADME properties, and so forth. In this approach it is easy to get trapped in local minima. Before the rise of target-based research, it was more common to drive programs based on animal efficacy data. This was, in effect, a form of simultaneous multifactorial optimization, since compounds that work in animal models are much more likely to avoid multiple problems, e.g., insolubility, cell impermeability, or metabolic instability.

Misuse of Structure-Based Drug Design. Structural information, predominantly from X-ray crystallography, frequently reveals obvious opportunities to fill deep pockets. A misguided but common design strategy focuses only on potency and will automatically increase molecular weight and, even worse, generally increase lipophilicity, since these pockets are usually hydrophobic.

However, it should not be assumed that structure-based design is more responsible for property inflation than other approaches. A 2000 publication by Lipinski[26] showed increases in a number of properties including molecular weight and lipophilicity for clinical candidates arising from both "rational" and HTS driven drug discovery programs. It would be enlightening to see an updated version of Lipinski's analysis and examine whether these trends have changed in the intervening years.

Dependence on High-Throughput Screening Collections (Rather than Natural Products) for Leads. In the mid-1990s, the HTS paradigm required building enormous chemical libraries, either through purchasing molecules from vendors or from high-throughput chemistry. Lead molecules coming from high-throughput screening often have poor physical properties, complicating the optimization process.[76]

Higher Hurdles to Compound Advancement. Over time, the requirements for high potency, selectivity, and other parameters have increased, which drives up molecular complexity and frequently reduces "druglikeness".

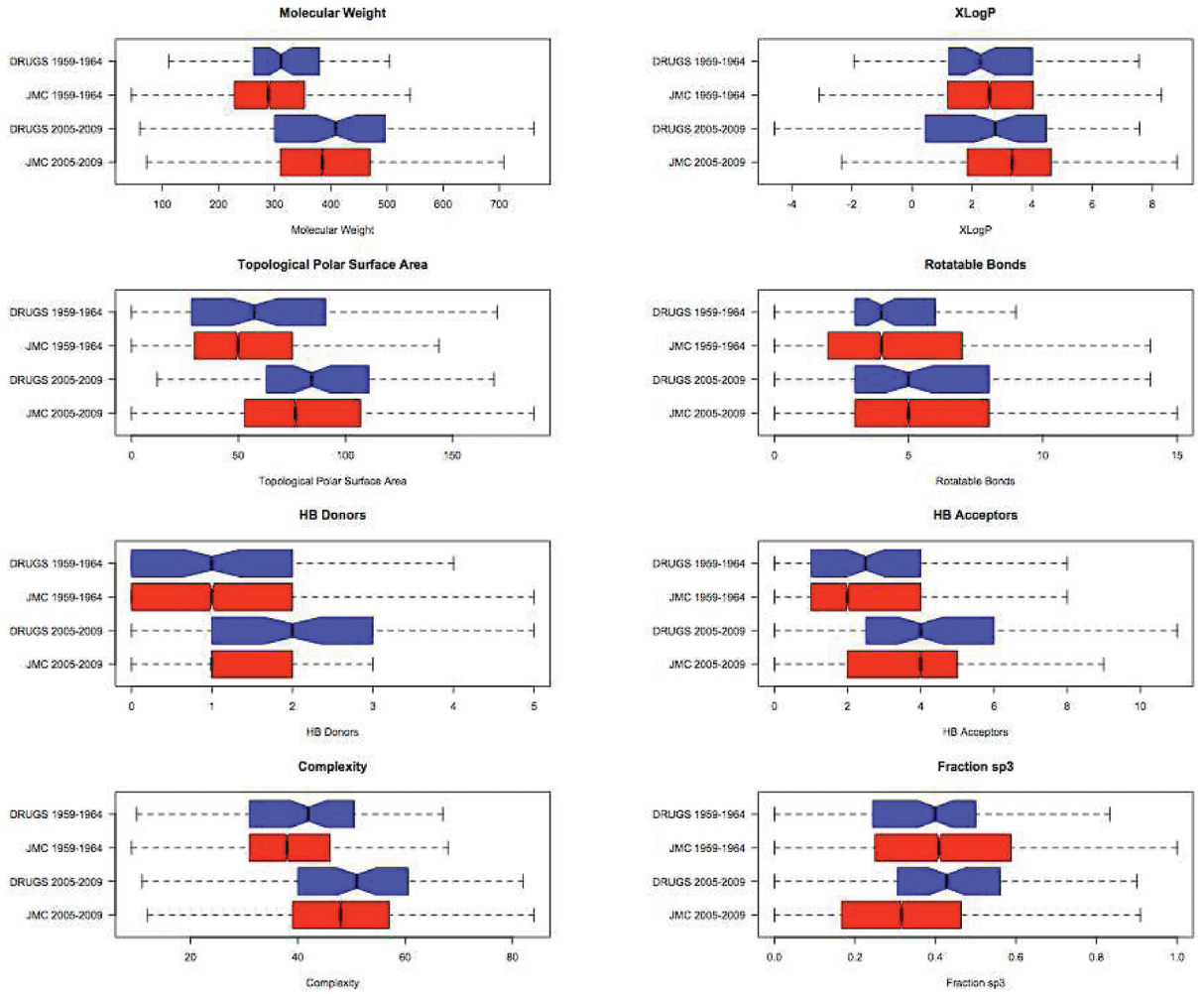

Figure 6. Comparison of property distributions for JMC compounds (in red) and launched drugs (in blue) for the first 5 years of JMC publication (1959−1964) with the most recent 5 years (2005−2009).

Changes in the Nature of JMC Manuscripts and/or Publication Policies. For completeness, we also include the possibility that these changes may also result from changes in JMC editorial policies. The fact that the increase in the number of molecules published in JMC has significantly outpaced the number of new drugs may indicate that earlier-stage projects, with less well-optimized compounds, are more likely to be reported in JMC in recent years.

CONCLUSIONS

If the *Journal of Medicinal Chemistry* is a fair surrogate for "what pharmaceutical chemists make", then we have confirmed the following facts: The mean and median values of a wide range of physical properties of molecules have risen, either steadily or in periodic jumps, throughout the past 50 years. Figures 2 and 3 summarize the properties of molecules published in JMC between 1959 and 2009. Both lipophilicity (as measured by XLogP) and three-dimensionality (as measured by Fsp3) are moving away from ranges typically found in successful drugs. There seems to be an acceleration in the upward trends of several properties, noticeably molecular weight and log *P*, reported in the journal around 1985. This provides a valuable clue: obviously *property changes were not solely due to high-throughput chemistry and screening*, since those technologies did not become commonplace until the mid-1990s. Likewise, the increase in molecular complexity is not a recent phenomenon and therefore cannot be solely attributed to factors such as combinatorial chemistry and high-throughput screening.

It is worth noting that these trends seemed to accelerate in the mid-1980s, indicating that some change took place in the early 1980s. The most likely explanations for an upward change in the early 1980s (before the age of combinatorial chemistry or high-throughput screening) seem to be advances in molecular biology, i.e., understanding of receptor subtypes leading to concerns about specificity; target-focused drug design and its corresponding one-property-at-a-time optimization paradigm (possibly exacerbated by structural biology); and improvements in technologies which enabled the synthesis and characterization of more complex molecules.

In recent years there has been a renewed interest in natural products and similar, more complex molecules as starting points for drug discovery programs.[77,78] As pointed out by Keller,[79] many natural products violate the rule of 5 but are still bioavailable. The reasons for these exceptions are not well understood, but further study should allow us to define valuable new rules.

Despite a perceived emphasis on "druglikeness", pharmaceutical output continues to decline. As of October 2009, drug approvals were down 29% over the previous year and 13% below the 10-year average.[11] If the compounds reported in JMC are, as we contend, a fair reflection of the sort of molecules being produced in drug discovery programs, it appears that much of what we have learned about "druglikeness" is not consistently being put into practice. To say this more precisely, despite the many publications (e.g., refs 9−16 and possibly others) outlining the challenges of developing compounds with properties that lie outside well-defined ranges, we find that the molecules actually being made by medicinal chemists are tending, more and more, to be such outliers, for the many good and understandable reasons described above.

Looking ahead, we believe it is necessary to select a different set of unifying principles for compound design. We will increasingly be driven to work on targets of unprecedented complexity while facing ever-increasing safety and efficacy hurdles. Simplistic rules,[9,10] even if widely adopted, will not suffice. A return to the practice, so widespread in the 1950s and 1960s, of relying solely on animal pharmacology as a primary readout seems unlikely. Rather, we propose five interdependent principles to improve the design process:

(1) Deeper understanding of true druglikeness. Simple numerical rules of thumb are helpful only to a degree. Rather, we need to develop an improved understanding of what makes some molecules druglike despite having physical properties *outside* the usual ranges. These rules can be global (applicable to all molecules) or local (fine-tuned for each chemical class or target tissue) and will be especially valuable for tackling "undruggable" targets.

(2) More emphasis on ligand efficiency,[28−30] particularly in the hit-to-lead process. This in turn would support a re-evaluation of screening collections and perhaps the use of fragment-based approaches. Greater ligand efficiency would also tend to lead to lower log *P*, which has clearly been shown[15] to reduce toxicological liabilities.

(3) More informative and higher throughput in vitro ADME/toxicity assays. Advances in assay technologies have led drug discovery teams to increasingly employ additional in vitro assays. Compounds in drug discovery programs are now routinely tested in solubility,[80] CYP,[81] and hERG assays.[82] In vitro systems such as Caco-2[83] and parallel artificial membrane permeation[84] are commonly used as surrogates for permeability. A number of cellular systems have been developed to provide an indication of potential adverse outcomes.[85,86] While many in vitro assays provide benefit, the correlation with in vivo data can be inconsistent. Increases in the quality and applicability of in vitro assays will enable medicinal chemists to gain insights earlier and avoid potential liabilities.

(4) A return to the mindset of simultaneous multivariate optimization. The ultimate goal of a drug discovery program is to generate a compound that is safe and efficacious in humans. In order to develop such a compound, teams must optimize many criteria including affinity, selectivity, activity, properties, and pharmacokinetics. These criteria are often optimized in a serial fashion. Teams will optimize a single criterion like enzyme potency and then sacrifice these gains to optimize a second property like solubility. The process of trading off one property for another can repeat for dozens of cycles. A number of approaches can be applied to support multiobjective optimization. One is to employ visualization software that enables teams to appreciate the entirety of the data. A number of groups have recently developed visualization tools oriented toward drug discovery programs.[87−89] Another approach is to apply multiobjective optimization algorithms and attempt to simultaneously optimize multiple criteria.[90,91] While multiobjective optimization has been successfully applied in a number of other fields, the primary limitation in drug discovery is the limited accuracy of computational models.[92,93]

(5) Reduced reliance on "easy" chemistry. While Pd-mediated sp^2-sp^2 couplings and amide bond-forming reactions have their place, we believe that a greater emphasis on the art of synthesis in medicinal chemistry would dramatically improve the physical properties of our molecules.

Adopting these principles will be challenging, but we believe that a significant change in approach is essential if medicinal chemistry is to remain a vibrant and productive enterprise.

■ ASSOCIATED CONTENT

Supporting Information. Box plots of calculated properties and mean and median property values for compounds published in *Journal of Medicinal Chemistry* between 1959 and 2009. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Voice: 617-444-6242. Fax: 617-444-6688. E-mail: pat_walters@vrtx.com.

■ BIOGRAPHIES

W. Patrick Walters is a Principal Research Fellow at Vertex Pharmaceuticals in Cambridge, MA, where he has worked since 1995. He heads the Computational Drug Discovery Technologies group, whose research focuses on the development of novel computational tools to support drug discovery and development. Before joining Vertex, Dr. Walters earned his Ph.D. in Organic Chemistry from the University of Arizona where he studied the application of artificial intelligence in conformational analysis. Prior to obtaining his Ph.D., he worked at Varian Instruments as both a chemist and a software developer. Dr. Walters received his B.S. in Chemistry from the University of California, Santa Barbara.

Jeremy Green obtained his B.Sc. from UMIST, Manchester, U.K., in 1984 and his Ph.D. in 1987 from the University of Edinburgh, U.K., working with Prof. Robert Ramage in the field of peptide chemistry. He was a Postdoctoral Fellow at the University of Wisconsin—Madison in the group of Prof. Daniel H. Rich working on HIV protease inhibition. Subsequently he joined Marion Merrell Dow (Indianapolis, IN), where he worked in early chemical development. He later joined Ariad Pharmaceuticals (Cambridge, MA) as a medicinal chemist. He has been at Vertex Pharmaceuticals since 1996 and is currently Director, Chemical Research.

Jonathan R. Weiss graduated from Trinity College, Hartford, CT, in 2007 with a B.S. in Chemistry. He subsequently joined the Discovery Chemistry group at Bristol-Myers Squibb working on

antiviral drug discovery. Since 2009, he has been working in the Computational Drug Discovery Technologies group at Vertex Pharmaceuticals.

Mark A. Murcko received his Ph.D. in physical organic chemistry from Yale University, CT, in 1987, working in the labs of Ken Wiberg. From 1987 to 1990 he worked at Merck Sharpe & Dohme in West Point, PA. Since 1990 he has been at Vertex Pharmaceuticals, where he now serves as Chief Technology Officer. In that role, his mission is to identify, evaluate, and bring into the organization a wide range of disruptive technologies that have the potential to transform R&D. He has authored more than 70 papers and is an inventor on more than 40 patents. He has contributed to four marketed drugs: dorzolamide, a carbonic anhydrase inhibitor for glaucoma; amprenavir and fosamprenavir, both HIV-PR inhibitors for AIDS; and telaprevir, an HCV-PR inhibitor for hepatitis C.

■ ACKNOWLEDGMENT

We thank our colleagues Youssef Bennani, Mike DeNinno, Jim Empfield, Peter Grootenhuis, Mike Mortimore, John Saunders, Dean Stamos, and Steve Young for their constructive comments on the manuscript. We also thank the anonymous reviewers for a number of useful suggestions. One reviewer's comment so nicely summarizes this analysis that we have chosen to directly include it: "Drugs have to survive multiple hurdles (the output of preclinical research is the start) followed by attritional factors including toxicity, clinical safety, efficacy in humans, differentiation, market viability, organizational strategy, regulatory approval, and acceptance by payers. It is not a surprise that "druglikeness" resists accurate description. However, the fact that one key property, lipophilicity, is unchanged in marketed drugs over time (but increasing in drug discovery) suggests that control of this property will remain essential."

■ ABBREVIATIONS USED

REOS, rapid elimination of swill; JMC, *Journal of Medicinal Chemistry*; CLogP, calculated octanol/water partition coefficient; EMBL, European Molecular Biology Laboratory; MDDR, MDL Drug Data Report; SMARTS, SMILES arbitrary target language; BMS, Bristol-Myers Squibb; PSA, polar surface area; hERG, human ether-a-go-go-related gene; BBB, blood−brain barrier

■ REFERENCES

(1) Almog, D. M. Drug industry in "depression". *Med. Sci. Monit.* **2005**, *11*, SR1–SR4.

(2) Booth, B.; Zemmel, R. Prospects for productivity. *Nat. Rev. Drug Discovery* **2004**, *3*, 451–456.

(3) Sams-Dodd, F. Target-based drug discovery: Is something wrong? *Drug Discovery Today* **2005**, *10*, 139–147.

(4) Schneider, C. K.; Schaffner-Dallmann, G. Typical pitfalls in applications for marketing authorization of biotechnological products in Europe. *Nat. Rev. Drug Discovery* **2008**, *7*, 893–899.

(5) Wood, A. J. A proposal for radical changes in the drug-approval process. *N. Engl. J. Med.* **2006**, *355*, 618–623.

(6) Owens, J. Big pharma slims down to bolster productivity. *Nat. Rev. Drug Discovery* **2007**, *6*, 173–174.

(7) Cuatrecasas, P. Drug discovery in jeopardy. *J. Clin. Invest.* **2006**, *116*, 2837–2842.

(8) Carney, S. How can we avoid the productivity gap? *Drug Discovery Today* **2005**, *10*, 1011–1013.

(9) Lipinski, C.; Lombardo, F.; Dominy, B.; Feeney, P. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.

(10) Walters, W. P.; Murcko, M. A. Prediction of "drug-likeness". *Adv. Drug Delivery Rev.* **2002**, *54*, 255–271.

(11) *Journal of Medicinal Chemistry*. http://pubs.acs.org/page/jmcmar/about.html.

(12) Lajiness, M. S.; Vieth, M.; Erickson, J. Molecular properties that influence oral drug-like behavior. *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 470–477.

(13) dVieth, M.; Siegel, M.; Higgs, R.; Watson, I.; Robertson, D.; Savin, K.; Durst, G.; Hipskind, P. Characteristic physical properties and structural fragments of marketed oral drugs. *J. Med. Chem.* **2004**, *47*, 224–232.

(14) Leeson, P.; Davis, A. Time-related differences in the physical property profiles of oral drugs. *J. Med. Chem.* **2004**, *47*, 6338–6348.

(15) 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.

(16) Leeson, P. D.; St-Gallay, S. A.; Wenlock, M. C. Impact of ion class and time on oral drug molecular properties. *Med. Chem. Commun.* **2011**, *2*, 91–105.

(17) Bender, A. Databases: compound bioactivities go public. *Nat. Chem. Biol.* **2010**, *6*, 309.

(18) De Matos, P.; Alcantara, R.; Dekker, A.; Ennis, M.; Hastings, J.; Haug, K.; Spiteri, I.; Turner, S.; Steinbeck, C. Chemical entities of biological interest: an update. *Nucleic Acids Res.* **2009**, D249–D254.

(19) Warr, W. A. ChEMBL. An interview with John Overington, team leader, chemogenomics at the European Bioinformatics Institute Outstation of the European Molecular Biology Laboratory (EMBL-EBI). *J. Comput.-Aided Drug Des.* **2009**, *23*, 195–198.

(20) Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714–3717.

(21) Wang, R.; Fu, Y.; Lai, L. A new atom-additive method for calculating partition coefficients. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 615–621.

(22) Wang, R.; Gao, Y.; Lai, L. Calculating partition coefficient by atom-additive method. *Perspect. Drug Discovery Des.* **2000**, *19*, 47–66.

(23) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1273–1280.

(24) Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.

(25) Proudfoot, J. R. The evolution of synthetic oral drug properties. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1087–1090.

(26) Lipinski, C. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.

(27) Gleeson, M. P. Generation of a set of simple, interpretable ADMET rules of thumb. *J. Med. Chem.* **2008**, *51*, 817–834.

(28) Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. Ligand binding efficiency: trends, physical basis, and implications. *J. Med. Chem.* **2008**, *51*, 2432–2438.

(29) Abad-Zapatero, C.; Metz, J. T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discovery Today* **2005**, *10*, 464–469.

(30) Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H. Ligand efficiency and fragment-based drug discovery. *Drug Discovery Today* **2009**, *14*, 278–283.

(31) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discovery* **2004**, *3*, 660–672.

(32) Hopkins, A. L.; Mason, J. S.; Overington, J. P. Can we rationally design promiscuous drugs? *Curr. Opin. Struct. Biol.* **2006**, *16*, 127–136.

(33) Schuffenhauer, A.; Brown, N.; Selzer, P.; Ertl, P.; Jacoby, E. Relationships between molecular complexity, biological activity, and structural diversity. *J. Chem. Inf. Model.* **2006**, *46*, 525–535.

(34) Posy, S. L.; Hermsmeier, M. A.; Vaccaro, W.; Ott, K. H.; Todderud, G.; Lippy, J. S.; Trainor, G. L.; Loughney, D. A.; Johnson,

S. R. Trends in kinase selectivity: insights for target class-focused library screening. *J. Med. Chem.* **2010**, *54*, 54–66.
(35) Waring, M. J. Lipophilicity in drug discovery. *Expert Opin. Drug Discovery* **2010**, *5*, 235–248.
(36) Martin, Y. C. A practitioner's perspective of the role of quantitative structure−activity analysis in medicinal chemistry. *J. Med. Chem.* **1981**, *24*, 229–237.
(37) Delaney, J. S. ESOL: estimating aqueous solubility directly from molecular structure. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1000–1005.
(38) Mackay, D.; Bobra, A.; Shiu, W. Y.; Yalkowsky, S. H. Relationships between aqueous solubility and octanol−water partition coefficients. *Chemosphere* **1980**, *9*, 701–711.
(39) Ran, Y.; Yalkowsky, S. H. Prediction of drug solubility by the general solubility equation (GSE). *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 354–357.
(40) Clark, D. E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. *J. Pharm. Sci.* **1999**, *88*, 807–814.
(41) Egan, W.; Merz, K., Jr.; Baldwin, J. Prediction of drug absorption using multivariate statistics. *J. Med. Chem.* **2000**, *43*, 3867–3877.
(42) Aronov, A. M. Common pharmacophores for uncharged human ether-a-go-go-related gene (hERG) blockers. *J. Med. Chem.* **2006**, *49*, 6917–6921.
(43) Jamieson, C.; Moir, E. M.; Rankovic, Z.; Wishart, G. Medicinal chemistry of hERG optimizations: highlights and hang-ups. *J. Med. Chem.* **2006**, *49*, 5029–5046.
(44) Ajay; Bemis, G. W.; Murcko, M. A. Designing libraries with CNS activity. *J. Med. Chem.* **1999**, *42*, 4942–4951.
(45) Clark, D. E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood−brain barrier penetration. *J. Pharm. Sci.* **1999**, *88*, 815–821.
(46) Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; DeCrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4872–4875.
(47) Kelder, J.; Grootenhuis, P. D. J.; Bayada, D. M.; Delbressine, L. P. C.; Ploemen, J. P. Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharm. Res.* **1999**, *16*, 1514–1519.
(48) Palm, K.; Stenberg, P.; Luthman, K.; Artursson, P. Polar molecular surface properties predict the intestinal absorption of drugs in humans. *Pharm. Res.* **1997**, *14*, 568–571.
(49) van de Waterbeemd, H.; Smith, D.; Beaumont, K. Property-based design: optimization of drug absorption and pharmacokinetics. *J. Med. Chem.* **2001**, *44*, 1313–1333.
(50) Hitchcock, S.; Pennington, L.; Lin, L.; Lanza, T., Jr. Structure −brain exposure relationships. *J. Med. Chem.* **2006**, *49*, 7559–7583.
(51) Veber, D.; Johnson, S.; Cheng, H.; Smith, B.; Ward, K. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623.
(52) Lu, J. J.; Crimin, K.; Goodwin, J. T.; Crivori, P.; Orrenius, C.; Xing, L.; Tandler, P. J.; Vidmar, T. J.; Amore, B. M.; Wilson, A. G. E.; Stouten, P. F. W.; Burton, P. S. Influence of molecular flexibility and polar surface area metrics on oral bioavailability in the rat. *J. Med. Chem.* **2004**, *47*, 6104–6107.
(53) Hou, T.; Wang, J.; Zhang, W.; Xu, X. ADME evaluation in drug discovery. 6. Can oral bioavailability in humans be effectively predicted by simple molecular-property based rules. *J. Chem. Inf. Model.* **2007**, *47*, 460–463.
(54) Blake, J. F. Examination of the computed molecular properties of compounds selected for clinical development. *BioTechniques* **2003**, No. Suppl., 16–20.
(55) Oprea, T. I. Property distribution of drug-related chemical databases. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 251–264.
(56) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J. Comb. Chem.* **1999**, *1*, 55–68.
(57) Goldstein, D. M.; Gray, N. S.; Zarrinkar, P. P. High-throughput kinase profiling as a platform for drug discovery. *Nat. Rev. Drug Discovery* **2008**, *7*, 391–397.
(58) Pors, K.; Goldberg, F. W.; Leamon, C. P.; Rigby, A. C.; Snyder, S. A.; Falconer, R. A. The changing landscape of cancer drug discovery: a challenge to the medicinal chemist of tomorrow. *Drug Discovery Today* **2009**, *14*, 1045–1050.
(59) Wells, J. A.; McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein−protein interfaces. *Nature* **2007**, *450*, 1001–1009.
(60) Schuffenhauer, A.; Brown, N.; Selzer, P.; Ertl, P. Relationships between molecular complexity, biological activity, and structural diversity. *J. Chem. Inf. Model.* **2006**, *46*, 525–535.
(61) Yan, A.; Gasteiger, J. Prediction of aqueous solubility of organic compounds by topological descriptors. *QSAR Comb. Sci.* **2003**, *22*, 821–829.
(62) 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.
(63) 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.
(64) Bemis, G. W.; Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **1996**, *39*, 2887–2893.
(65) Bemis, G. W.; Murcko, M. A. Properties of known drugs. 2. Side chains. *J. Med. Chem.* **1999**, *42*, 5095–5099.
(66) 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.
(67) Heck, R. F.; Nolley, J. P., Jr. Palladium-catalyzed vinylic hydrogen substitution reactions with aryl, benzyl, and styryl halides. *J. Org. Chem.* **1972**, *37*, 2320–2322.
(68) King, A. O.; Okukado, N.; Negishi, E. Highly general stereo-, regio-, and chemo-selective synthesis of terminal and internal conjugated enynes by the Pd-catalysed reaction of alkynylzinc reagents with alkenyl halides. *J. Chem. Soc., Chem. Commun.* **1977**, 683–684.
(69) Chambers, J. M.; Cleveland, W. S.; Kleiner, B.; Tukey, P. A. In *Graphical Methods for Data Analysis*; Wadsworth & Brooks/Cole: Belmont, CA, 1983; p 62.
(70) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008**, *26*, 127–132.
(71) Fowler, P. W.; Tai, K.; Sansom, M. S. P. The selectivity of K^+ ion channels: testing the hypotheses. *Biophys. J.* **2008**, *95*, 5062–5072.
(72) Kenakin, T. Functional selectivity in GPCR modulator screening. *Comb. Chem. High Throughput Screening* **2008**, *11*, 337–343.
(73) Hawkins, M. J.; Soon-Shiong, P.; Desai, N. Protein nanoparticles as drug carriers in clinical medicine. *Adv. Drug Delivery Rev.* **2008**, *60*, 876–885.
(74) Li, J.; Loh, X. J. Cyclodextrin-based supramolecular architectures: syntheses, structures, and applications for drug and gene delivery. *Adv. Drug Delivery Rev.* **2008**, *60*, 1000–1017.
(75) Vintiloiu, A.; Leroux, J. C. Organogels and their use in drug delivery: a review. *J. Controlled Release* **2008**, *125*, 179–192.
(76) Keser, G. M.; Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discovery* **2009**, *8*, 203–212.
(77) Koehn, F. E.; Carter, G. T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discovery* **2005**, *4*, 206–220.
(78) Li, J. W. H.; Vederas, J. C. Drug discovery and natural products: end of an era or an endless frontier? *Science* **2009**, *325*, 161–165.
(79) Keller, T.; Pichota, A.; Yin, Z. A practical view of "druggability". *Curr. Opin. Chem. Biol.* **2006**, *10*, 357–361.
(80) Wyttenbach, N.; Alsenz, J.; Grassmann, O. Miniaturized assay for solubility and residual solid screening (SORESOS) in early drug development. *Pharm. Res.* **2007**, *24*, 888–898.

(81) Lahoz, A.; Donato, M. T.; Castell, J. V.; Gomez-Lechon, M. J. Strategies to in vitro assessment of major human CYP enzyme activities by using liquid chromatography tandem mass spectrometry. *Curr. Drug Metab.* **2008**, *9*, 12–19.

(82) Hoffmann, P.; Warner, B. Are hERG channel inhibition and QT interval prolongation all there is in drug-induced torsadogenesis? A review of emerging trends. *J. Pharmacol. Toxicol. Methods* **2006**, *53*, 87–105.

(83) Sambuy, Y.; De Angelis, I.; Ranaldi, G.; Scarino, M. L.; Stammati, A.; Zucco, F. The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics. *Cell Biol. Toxicol.* **2005**, *21*, 1–26.

(84) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J. Med. Chem.* **1998**, *41*, 1007–1010.

(85) Xu, J. J.; Henstock, P. V.; Dunn, M. C.; Smith, A. R.; Chabot, J. R.; De Graaf, D. Cellular imaging predictions of clinical drug-induced liver injury. *Toxicol. Sci.* **2008**, *105*, 97–105.

(86) Greene, N.; Aleo, M. D.; Louise-May, S.; Price, D. A.; Will, Y. Using an in vitro cytotoxicity assay to aid in compound selection for in vivo safety studies. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5308–5312.

(87) Kolpak, J.; Connolly, P. J.; Lobanov, V. S.; Agrafiotis, D. K. Enhanced SAR maps: expanding the data rendering capabilities of a popular medicinal chemistry tool. *J. Chem. Inf. Model.* **2009**, *49*, 2221–2230.

(88) Ritchie, T. J.; Ertl, P.; Lewis, R. The graphical representation of ADME-related molecule properties for medicinal chemists. *Drug Discovery Today* **2011**, *16*, 65–72.

(89) Wassermann, A. M.; Wawer, M.; Bajorath, J. Activity landscape representations for structure−activity relationship analysis. *J. Med. Chem.* **2010**, *53*, 4891–4896.

(90) Manly, C. J.; Chandrasekhar, J.; Ochterski, J. W.; Hammer, J. D.; Warfield, B. B. Strategies and tactics for optimizing the hit-to-lead process and beyond: a computational chemistry perspective. *Drug Discovery Today* **2008**, *13*, 99–109.

(91) Nicolaou, C. A.; Brown, N.; Pattichis, C. S. Molecular optimization using computational multi-objective methods. *Curr. Opin. Drug Discovery Dev.* **2007**, *10*, 316–324.

(92) Doweyko, A. M. QSAR: dead or alive? *J. Comput.-Aided Mol. Des.* **2008**, *22*, 81–89.

(93) Johnson, S. R. The trouble with QSAR (or how I learned to stop worrying and embrace fallacy). *J. Chem. Inf. Model.* **2008**, *48*, 25–26.