Time-Related Differences in the Physical Property Profiles of Oral Drugs

Paul D. Leeson*,† and Andrew M. Davis‡

Department of Medicinal Chemistry and Department of Physical and Metabolic Science, AstraZeneca R&D Charnwood, Bakewell Road, Loughborough, Leicestershire LE11 5RH, U.K.

Received April 13, 2004

Comparisons of the calculated physicochemical properties of oral drugs launched prior to 1983 (864 drugs) and between 1983 and 2002 (329 drugs) show that mean values of lipophilicity, percent polar surface area and H-bond donor count are the same, suggesting that these are the most important oral druglike physical properties. In contrast, mean values of molecular weight and the numbers of O + N atoms, H-bond acceptors, and rotatable bonds and rings have increased in 1983−2002 drugs (by 13−29%). Analysis of the 1983−2002 oral drugs by therapy area shows that antiinfectives and nervous system drugs have the most extreme physical property profiles. Cardiovascular drugs show increasing molecular weight with year of publication, primarily a consequence of focusing on clinically proven mechanisms, with limited chemical diversity. Drug classes other than antiinfectives show comparable distributions of lipophilicity, suggesting that this property in oral drugs is important irrespective of the drug's target. The results suggest that the balance between polar and nonpolar drug properties is an important, unchanging feature of oral drug molecules.

Introduction

Despite increased expenditure in research and development activities, the output of launched drugs has declined in recent years.[1] Only 11% of drugs entering clinical development reach the market place, being withdrawn for reasons associated with efficacy (25%), toxicology (24%), clinical safety (12%), drug metabolism and pharmacokinetics (DMPK, 8%), formulation (1%), and portfolio and other reasons (30%).[2] Thus, of 70% of failures caused by specific effects, 45% can be ascribed to DMPK, safety, and formulation properties related to the physicochemical nature, or druglikeness, of the drug candidate itself. The proportion may even be higher, since some reported "efficacy" failures might be due to poor DMPK.

While there have been numerous attempts to predict druglikeness,[3] currently no generally applicable means of differentiating drugs from nondrugs exists. Many of the methods proposed require complex computational analyses, leading some authors to suggest simpler alternatives based on chemical intuition and ease of synthesis.[4] The most widely quoted method, the Lipinski "rule of five", is also the simplest. This states that poor drug absorption and permeation are likely to occur if the molecular weight is >500, cLogP (or calculated lipophilicity, the logarithm of the 1-octanol/water partition coefficient) is >5, the sum of O and N atoms is >10, and the sum of OH and NH groups is >5.[5] The rule of five was designed to provide chemists with a simple means of predicting potential problems with solubility and permeability, factors strongly influencing drug absorption.[6] It has been adopted, probably inappropriately, as a rule of thumb for druglike properties in the broadest sense; however, drugs also have to be efficacious, potent, selective, manufacturable, and safe. The rule of five was derived from drug candidates that reached phase II (receiving a U.S. adopted name), so many of these compounds will have eventually failed to progress to the market. Consequently recent studies[7,8] of druglike physical properties have focused solely on the smaller number of marketed oral drugs, which have successfully passed DMPK, formulation, manufacturing, toxicological, and clinical hurdles in the drug development process.

The studies of Wenlock and co-workers[7] (594 compounds from the Physicians Desk Reference 1999 (PDR)) and Vieth and co-workers[8] (1193 compounds approved by the U.S. Food and Drug Administration up to 2002) reached essentially identical conclusions, showing that limited distributions of molecular weight, lipophilicity, and hydrogen bonding are found in oral drugs. It was shown in both studies that oral drug physical properties are in general consistent with the rule of five but with greater stringency for H-bonding properties. Thus, for 90% of oral drugs it was found that molecular weight was <475, cLogP was <5.2−5.5, the sum of O and N atoms was <8−9, and the sum of OH and NH groups was <3.[7,8] In addition, overall physical properties of oral drugs are reduced compared with topical, injectable, and absorbent drugs and early research compounds.[8] The observations are essentially consistent with the established dependencies of solubility, permeability, absorption, chemical and metabolic stability, and toxicity on the bulk and hydrogen-bonding properties of compounds.[9,10]

Vieth[8] showed that among oral drugs launched in the U.S. between 1982 and 2002, no meaningful correlations between year of launch and molecular weight, lipophilicity, or target class were observed, suggesting that there is an important and unchanging need to maintain physical properties within a particular range to allow

* To whom correspondence should be addressed. Phone: +44 (0)-1509 644826. Fax: +44 (0)1509 645571. E-mail: paul.leeson@astrazeneca.com.
† Department of Medicinal Chemistry.
‡ Department of Physical and Metabolic Science.

10.1021/jm049717d CCC: $27.50
Published on Web 11/02/2004

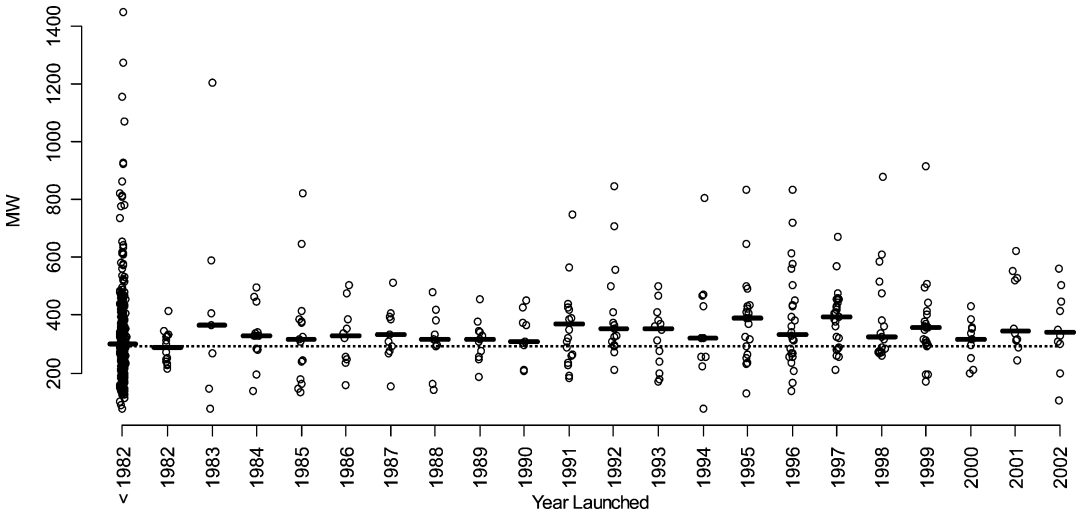

Figure 1. Molecular weight versus year of launch in the U.S. for oral drugs. The median molecular weights of drugs launched in each year from 1983 to 2002 are higher than those of the drugs launched prior to 1983. Adapted with permission from *J. Med. Chem*. **2004**, *47*, 224−232.[8]

sufficient permeability. Wenlock[7] analyzed drugs in various phases of clinical development and showed that the mean molecular weight of compounds in earlier development phases was significantly higher than those in later development phases and marketed drugs. Similar trends were found for the numbers of hydrogen-bond acceptors and rotatable bonds, and in addition, it was shown that the more lipophilic compounds tend to be discontinued prior to phase III. These observations, supported by other studies,[11] suggest that there may be pressures during drug development to select smaller and less lipophilic compounds. Alternatively, it has been noted that they could simply be a consequence of larger and more lipophilic candidate drugs entering development.[12]

Can the observed limitations[7,8] to the physical properties of historical drugs be challenged? One recent study,[13] based on GlaxoSmithKline compounds, suggests that the key factors controlling bioavailability are polar surface area and molecular flexibility (number of rotatable bonds), and it was suggested, perhaps controversially,[10] that drug molecular weight might be increased if these properties are kept under control.

In this paper we ask which are the most important physical properties of oral drugs and attempt to answer the question by comparing the physical properties of newer (1983−2002) and older (pre-1983) oral drugs. We propose that drug physical properties, which are not significantly different between these two groups, are those to which the greatest attention should be paid in the design process and in candidate drug selection; conversely those that are different are potentially less significant and may be open to wider manipulation. It has been recognized that physical properties of drugs can be influenced by the targeted therapeutic area,[14] and here, we compare the profiles of recent drugs, launched from 1983 to 2002, in five major therapy areas.

Data Sources

Although Vieth[8] found there was no significant year-on-year change in oral drug molecular weight between 1982 and 2002, close examination of the published data (reproduced in Figure 1) shows that those oral drugs approved in the U.S. for each of the years from 1983 to 2002 have higher median molecular weights than 1982 and pre-1982 drugs. In contrast, little or no change in lipophilicity among these drugs is apparent from the corresponding lipophilicity vs year of launch correlation.[8] To test the significance of these observations, we used the list of oral drugs and their physical properties from the supplementary data kindly provided by Vieth[8] (1193 compounds). On the basis of the data in Figure 1, we decided to choose the year 1983 as the cutoff point for "older" drugs. Those drugs which first reached the market from 1983 to 2002 were readily identified from the valuable resource compiled by the ACS publication *Annual Reports in Medicinal Chemistry*.[15] This provided both the year of first marketing anywhere in the world, rather than just the U.S. approval date as used in ref 8, and the therapeutic indication. This resulted in two sets of oral drugs: those first launched prior to 1983 (864 compounds) and those first launched from 1983 to 2002 (329 compounds). Physical properties examined were molecular weight, lipophilicity (estimated logarithm of the 1-octanol/water partition coefficient, cLogP), percent polar surface area ([(polar surface area)/(total surface area)] × 100, values based on a 3D method[8]), sum of OH + NH groups (H-bond donors), sum of O + N atoms, number of H-bond acceptors, number of rotatable bonds, and number of rings.

Results and Discussion

Physical Property Comparisons of Older (Pre-1983) and Newer (1983−2002) Oral Drugs. Mean and median values of molecular weight, sum of O + N atoms, H-bond acceptors, rotatable bonds, and number of rings are significantly increased in 1983−2002 vs pre-1983 oral drugs by 13−29% (Table 1, shown graphically in Figure 2). In contrast, mean values of lipophilicity (cLogP), percent polar surface area, and the sum of OH + NH groups are not different, suggesting that these properties are of more fundamental importance in oral drugs. The choice of 1983 as the cutoff does not appear to greatly influence these findings. A later cutoff year could have been used because there are no meaningful correlations between any physical property and year of

Table 1. Mean (Median) Physical Properties of Oral Drugs Launched Pre-1983 and 1983−2002[a]

	oral drugs pre-1983[b] $n = 864$	oral drugs 1983−2002[c] $n = 329$	p pre-1983 vs 1983−2002[d]	difference in mean (median) values
Mol Wt	331 (310)	377 (357)	5.82×10^{-7}	14% (15%)
cLogP	2.27 (2.31)	2.50 (2.36)	0.17	10% (2%)
%PSA	21.1 (18.5)[e]	21.0 (19.4)	0.90	0% (5%)
OH + NH	1.81 (1)	1.77 (1)	0.35	−2% (0%)
O + N	5.14 (4)	6.33 (6)	5.65×10^{-8}	23% (50%)
HBA	2.95 (2)	3.74 (3)	1.34×10^{-7}	27% (50%)
RotB	4.97 (4)	6.42 (6)	2.20×10^{-8}	29% (50%)
Rings	2.56 (3)	2.88 (3)	1.18×10^{-4}	13% (0%)

[a] Mol Wt = molecular weight; cLogP = calculated 1-octanol/water partition coefficient (Daylight method); %PSA = calculated [(polar surface area)/(total surface area)] × 100; OH + NH = sum of OH + NH groups (H-bond donors); O + N = sum of O + N atoms; HBA = sum of H-bond acceptors; RotB = number of freely rotating bonds; Rings = number of rings in structure. All physical property values are taken from ref 8. [b] U.S. drugs (FDA approved to 2002) from Supporting Information in ref 8, excluding 83-02 NCEs. [c] Identified from oral drugs in Supporting Information in ref 8, using the 1983−2002 NCE list in ref 15. [d] Two-tailed, from two-sample t-test assuming unequal variances. [e] $n = 860$ (four compounds have missing values in ref 8).

launch among the 1983−2002 oral drugs.[8] In addition, comparisons of oral drugs launched in the decades 1983−1992 and 1993−2002 show that while there are upward trends in most properties with time, none reach statistical significance (Table 2). Comparing the data in Tables 1 and 2 suggests that oral drugs launched most recently (1993−2002) have contributed to the increases in mean cLogP (though not significant) and number of rings seen in the 1983−2002 vs pre-1983 comparisons.

Drugs launched from 1983 to 2003 are on average 46 Da larger than pre-1983 drugs, a mean increase of 14%. The distributions in Figure 3 show that violations of the Lipinski rule of molecular weight (>500) comprise 11.3% of 1983−2002 drugs and 6.7% of pre-1983 drugs. Increased molecular weight is accompanied by no change in mean lipophilicity (Figure 4), so there must be an increase in polarity or H-bonding in 1983−2002 drugs. This seems to be confined to O + N atoms (and H-bond acceptors) rather than H-bond donors (Figure 5) or percent polar surface area (Figure 6). Consistent with increased molecular weight in 1983−2002 drugs is increased structural complexity, as adjudged by increased numbers of rotatable bonds and rings (Figure 7). There has been a 29% increase in mean number of rotatable bonds (Table 1) in 1983−2002 drugs, and moreover, the property distribution is not normal, with higher frequencies of >9 increased relative to pre-1983 drugs (Figure 7). These observations contrast with recent studies showing that lowered bioavailability is linked to increased number of rotatable bonds.[13] This study used a proprietary set of GlaxoSmithKline compounds, and so the extent to which the specific structural properties and chemical diversity of these compounds influenced the published analysis cannot be ascertained. While bioavailability among a subset of established drugs could also be categorized on the basis of rotatable bonds and polar surface area,[13] further studies exploring the correlation between bioavailability and rotatable bonds are warranted.

The physical properties used in Table 1 are simple to calculate, and their use has gained widespread acceptance, but it has been known for some time that the bulk physical properties of molecules are correlated.[16] Hence, one of the issues in using the physical properties in Table 1 is potential redundancy. This is illustrated simply among the full set of oral drugs, where the four "Lipinski parameters" are clearly linked:

$$\text{cLogP} = 0.018(\text{Mol Wt}) - 0.64(\text{O} + \text{N}) - 0.40(\text{OH} + \text{NH}) + 0.19 \quad (1)$$

$$n = 1193; \quad r = 0.79; \quad t_{\text{Mol Wt}} = 39.4; \quad t_{\text{O+N}} = 28.6; \quad t_{\text{OH+NH}} = 12.0$$

All the variables are highly statistically significant in eq 1, which is a very simplified version of more sophisticated regression analyses for exploring the underlying molecular properties influencing solution-phase partition, where molar volume, H-bonding basicity and acidity, and polarizability were shown to be important.[17] Since molecular weight and O + N count have increased in 1983−2002 drugs and cLogP has not, eq 1 seems to suggest that the single most important invariant molecular property of oral drugs is the number of H-bond donors. However, the distributions of cLogP and percent polar surface area values show that the *balance* of polar versus nonpolar properties is critical.

The increases in molecular size and structural complexity in newer drugs are clearly acceptable, but compounds reported in phase I (mean molecular weight of 424[7]) or in the research phase (mean molecular weight of 447[8]) remain notably larger than 1993−2002 drugs (mean molecular weight 382, Table 2). Higher molecular weight compounds, with a greater number of pharmacophoric groups, will have additional potential for metabolic transformations to other biologically active and/or chemically reactive molecules. Increased solute molecular weight also reduces rates of diffusion,[18] and because molecular weight is positively correlated with the other physical properties of oral drugs,[8] increasing it will tend to increase other properties also. In addition, high molecular weight (>550) is associated with hepatic clearance via biliary or metabolic excretion.[18] For these reasons it is improbable that molecular weight can continue to increase indefinitely in new oral drug molecules.

The cumulative fraction plot in Figure 4 reveals that there has been a small increase in lipophilicity but only in the lower 50% of compounds, resulting in a slightly narrower range in the 1983−2002 drug molecules. Although the lipophilicity distributions of newer and older drugs are the same, the range of allowable cLogP values is large, the 10−90 percentile values from Figure 4 are the following: 1983−2002 drugs, −0.65 − 5.36; pre-1983 drugs, −0.84 − 5.18. However, the lipophilicity requirements within a particular chemical class are likely to be much narrower, being determined by specific structure−activity, selectivity, and metabolic and pharmacokinetic properties. The essentially unchanging overall lipophilicity distribution in oral drugs reflects the widely accepted importance of this property in binding to protein targets and in drug transport, usually resulting in a need for optimal values of lipophilicity for in vivo activity.[19,20] There are ∼1700 published quantitative structure−activity relationships on a large range of biological effects from binding affinity to

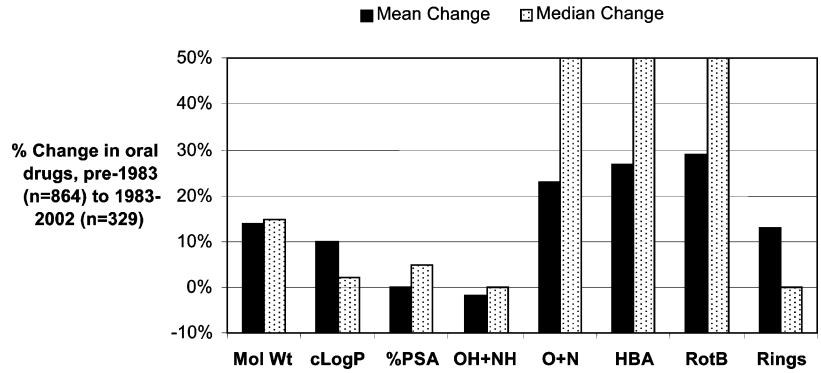

Figure 2. Mean and median differences in physical properties between pre-1983 and 1983−2002 drugs (data from Table 1): Mol Wt = molecular weight; cLogP = calculated 1-octanol/water partition coefficient (Daylight method); %PSA = calculated [(polar surface area)/(total surface area)] ×100; OH + NH = sum of OH + NH groups (H-bond donors); O + N = sum of O + N atoms; HBA = sum of H-bond acceptors; RotB = number of freely rotating bonds; Rings = number of rings in structure.

Table 2. Mean (Median) Physical Properties of Oral Drugs Launched 1983−1992 and 1993−2002[a]

	oral drugs 1983−1992[b] $n = 175$	oral drugs 1993−2002[b] $n = 154$	p 1983−1992 vs 1993−2002[c]	difference in mean (median) values
Mol Wt	374 (359)	382 (357)	0.62	2.1% (−0.6%)
cLogP	2.39 (2.36)	2.61 (2.38)	0.41	9.2% (0.8%)
%PSA	20.9 (19.0)	21.2 (19.7)	0.81	1.4% (3.7%)
OH + NH	1.75 (1)	1.80 (1.5)	0.76	2.9% (50%)
O + N	6.33 (6)	6.32 (6)	0.97	0.2% (0%)
HBA	3.66 (3)	3.82 (4)	0.51	4.4% (33%)
RotB	6.29 (6)	6.58 (6)	0.51	4.6% (0%)
Rings	2.77 (3)	3.02 (3)	0.071	9.0% (0%)

[a] Mol Wt = molecular weight; cLogP = calculated 1-octanol/water partition coefficient (Daylight method); %PSA = calculated [(polar surface area)/(total surface area)] × 100; OH + NH = sum of OH + NH groups (H-bond donors); O + N = sum of O + N atoms; HBA = sum of H-bond acceptors; RotB = number of freely rotating bonds; Rings = number of rings in structure. All physical property values are taken from ref 8. [b] Compounds from 1983 to 2002 drugs (Table 1). [c] Two-tailed, from two-sample t-test assuming unequal variances.

toxicity, showing that increasing lipophilicity (either cLogP of the whole molecule or π values of a substituent) increases biological activity.[19] As lipophilicity (cLogP) increases, there is an increased probability of binding to hydrophobic protein targets other than the desired one, and therefore, there is more potential for toxicity. For example, the active sites for undesirable cytochrome P450 inhibition[21] and the hERG channel[22] are hydrophobic and bind lipophilic substrates. Increasing lipophilicity will also tend to reduce aqueous solubility, reduce the free fraction of drugs by increased binding to plasma proteins, and increase metabolic liability.[9,10]

The lack of differences in mean values of H-bond donors and percent polar surface area between newer and older oral drugs may be related to an underlying need for effective membrane permeability. The results suggest that hydrogen-bond donors may be more important than acceptors in this respect, an interpretation that deserves comment in the context of current models of membrane permeability. Membrane permeability requires both desolvation of associated hydrogen-bonded water molecules as well as lipid solubility, and it is known that cell penetration and absorption are highly dependent on hydrogen bonding as well as size and lipophilicity.[23,24] 1-Octanol/water partition coefficients (ie $\log P$ values) alone appear insufficient to model membrane permeation. Work reported by Abraham and co-workers suggests that 1-octanol as a hydrophobic solvent is a poor model of the hydrophobic core of the membrane phospholipid bilayer. Water is a much stronger hydrogen-bond acid than 1-octanol, but both have similar hydrogen-bond basicities. This indicates that partitioning into 1-octanol underestimates the energy required to desolvate solute hydrogen-bond donors[25] for partition into a biological membrane. Thus, optimization of drug molecules using n-octanol/water partition coefficients requires additional consideration of hydrogen-bond donor properties.

Although mean and median values do not differ, it is notable that the distribution of percent polar surface area in 1983−2002 drugs is significantly narrower than pre-1983 drugs (see Figure 6). The 10−90 percentile values of percent polar surface area, 4.5−39.5% for pre-1983 drugs and 9.9−35.9% for 1983−2002 drugs, show that the distribution has narrowed by 25% in 1983−2002 drugs. Similar distributions were found using percent O + N atoms instead of percent polar surface area (data not shown: O + N atom count and polar surface area are highly correlated[8]). The reasons for the narrower distributions of percent polar surface area and percent O + N atoms may be related to the increased size and structural complexity of the 1983−2002 drugs, necessitating a more restricted, better-balanced polar/nonpolar profile. It could also result from the deliberate optimization of compound physical properties during synthesis−testing cycles, which will have been employed more rigorously in recent drug discovery.

Most of the oral drugs approved from 1983 to 2002 would have been discovered from the mid-1970s to the 1990s. Drug discovery projects during this time (and today) in general used in vitro screens as a primary test, a change from much earlier approaches, which often depended on in vivo animal studies to a greater extent in initial evaluation. In addition, newer drug discovery programs in general undertake more intensive and faster biological evaluation of compounds, facilitated by using rapid synthesis and testing techniques. The increases in molecular weight and other properties seen in the comparison of pre-1983 and post-1983 drugs are most probably a result of focusing on enhancing the affinity of lead compounds for the biological target in vitro, which is often achieved by adding interactive H-bonding or hydrophobic groups. Furthermore, commonly used strategies to improve solubility, for example,

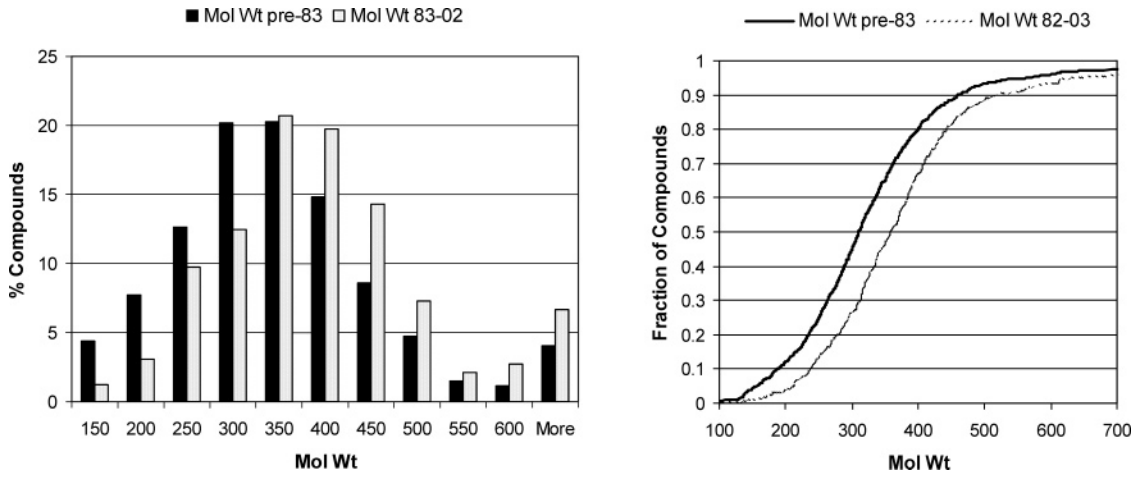

Figure 3. Distribution of molecular weight in pre-1983 (n = 864) and 1983−2002 (n = 329) oral drugs: (left) frequency; (right) cumulative fraction.

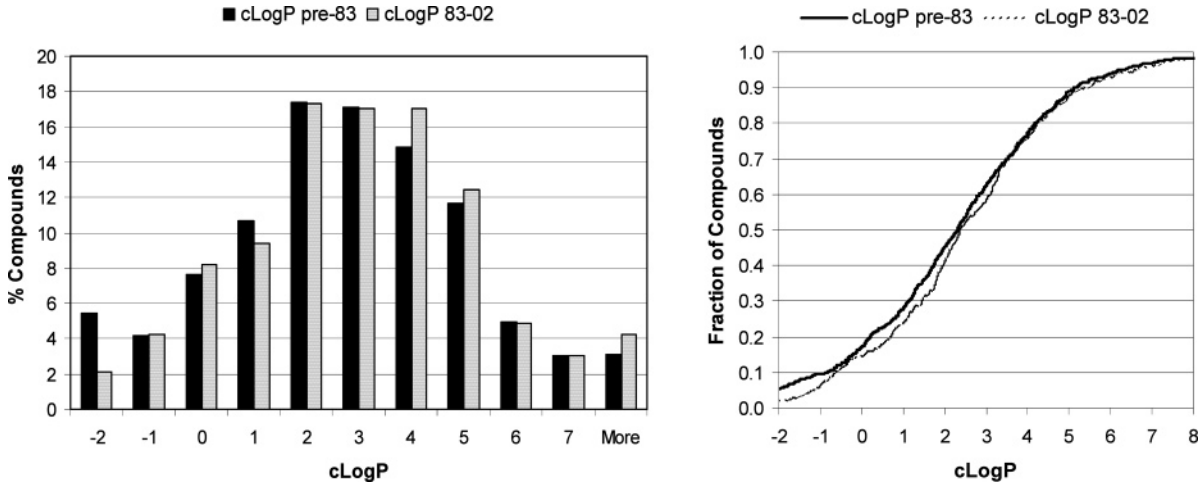

Figure 4. Distribution of lipophilicity (Daylight cLogP) in pre-1983 (n = 864) and 1983−2002 (n = 329) oral drugs: (left) frequency; (right) cumulative fraction. The differences in the lower 50% mean values are significant: pre-1983, 0.26; 1983−2002, 0.67; p = 0.0055 (two tailed, from two-sample t-test assuming unequal variances).

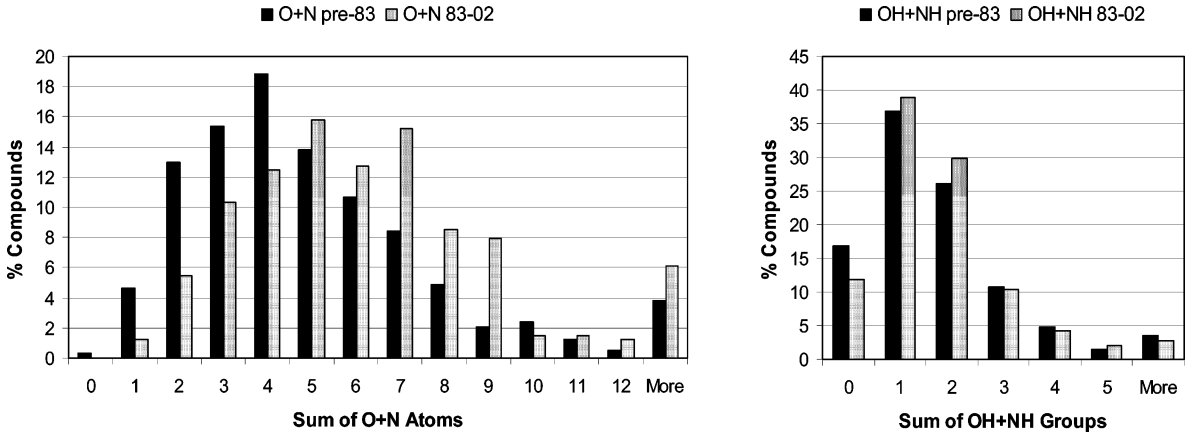

Figure 5. Frequency distributions of O + N atoms (left) and OH + NH groups (right) in pre-1983 (n = 864) and 1983−2002 (n = 329) oral drugs.

by adding ionizable functionality, and selectivity for the desired target, by adding groups to block unwanted activity, can result in increased molecular weight. It has been suggested[26] that the process of optimization of lead compounds in general results in increased molecular weight and other physical properties. This phenomenon is becoming more widely recognized,[27−29] supporting the proposal that optimization should ideally begin with identifying small, low molecular weight or "lead-like"[26] compounds.

Other reasons underlying the increases in some physical properties in newer drug candidates may be associated with medicinal strategies used to identify lead compounds. In a large number of Pfizer early drug candidates, molecular weight and lipophilicity increased with time but O + N atom counts did not.[30] In contrast,

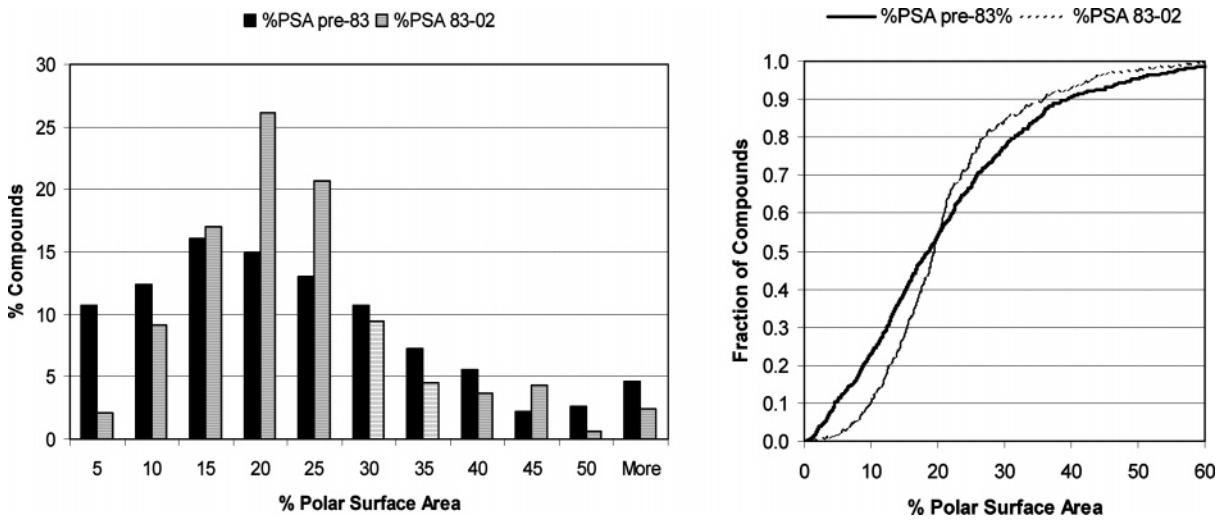

Figure 6. Distribution of percent polar surface area in pre-1983 ($n = 860$) and 1983−2002 ($n = 329$) oral drugs: (left) frequency; (right) cumulative fraction. The mean values of the lower and upper 50% of compounds are the following: pre-1983, 10.2% and 32.0%; 1983−2002, 13.5% and 28%. The differences are significant: p(lower 50%) = 4.1×10^{-14}, p(upper 50%) = 3.3×10^{-4} (two tailed, from two-sample t-test assuming unequal variances).

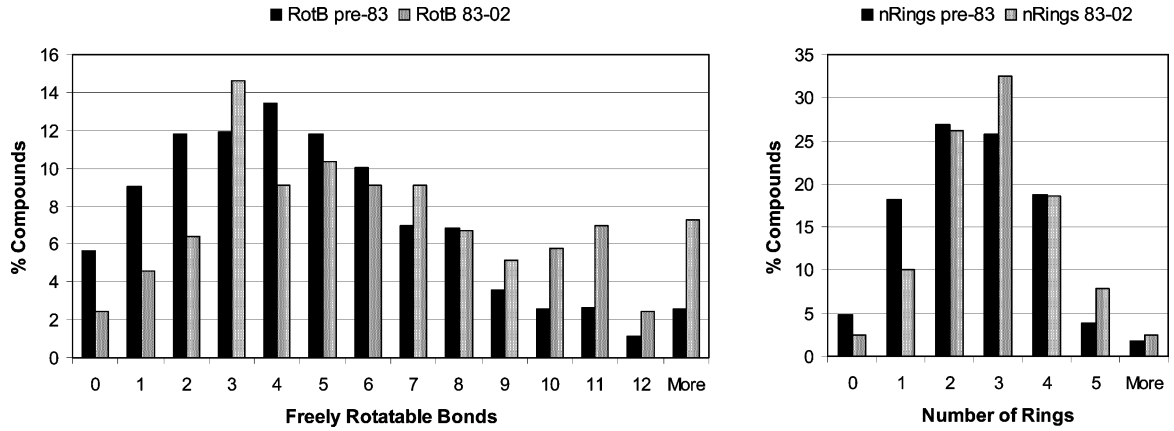

Figure 7. Frequency distributions of numbers of freely rotatable bonds (left) and rings (right) in pre-1983 ($n = 864$) and 1983−2002 ($n = 329$) oral drugs.

Merck MK-numbered development candidates, considered to be more advanced in development than the Pfizer candidates,[30] had increased molecular weight and O + N atom count with time but no change in mean lipophilicity. These differences were suggested to be a result of differing approaches to lead generation: structure-based and rational discovery for Merck and high-throughput screening with combinatorial chemistry (providing higher molecular weight and more lipophilic lead compounds) for Pfizer. Comparing these findings with the oral drug profiles at different development stages[7] suggests that the stage of development reached may be an additional factor. Lower mean lipophilicity is evident in the more advanced Merck candidates (cLogP ≈ 2.2) compared with the earlier Pfizer candidates (cLogP = 2−3), where development attrition will not have occurred to the same extent.

Influence of Therapeutic Area on 1983−2002 Oral Drug Physical Properties. Drug physical properties can be influenced by the targeted therapeutic area, for example, drugs aimed at the central nervous system tend to be smaller and have lower polar surface area than other classes.[14,17,31] On the basis of the literature classification,[15] the therapy areas of 315 of the 329 oral drugs launched from 1983 to 2002 were grouped into six categories (Table 3). Mean and median values of the eight physical properties are shown in Table 4, and the relative mean values are plotted in Figure 8 for the five major therapeutic categories. The results of pairwise comparisons, checking for statistically meaningful differences, are summarized in Table 5 for the five largest therapeutic categories.

The variance in physical property profile across therapeutic areas is significant (Figure 8) with a total of 49 of the 80 possible pairwise comparisons between classes being different at a 95% confidence level ($p < 0.05$, Table 5). This is probably a consequence of the relatively small numbers of drugs, and especially limited structural diversity, found in each therapeutic category. Comparing the therapeutic class properties at higher statistical confidence levels shows that a total of 21 of the 22 differences observed at a confidence level of >99.9% ($p < 0.001$; these properties are in bold in Table 5) are associated with either antiinfective or nervous system drug properties. Similarly, 10 of the 14 differences, significant at the 99−99.9% confidence level ($p = 0.001-0.01$; these properties are in italic in Table 5), are also associated with the antiinfectives and nervous system drugs, which clearly show the greatest overall differences in physical properties from the other classes.

Table 3. Major Therapy Areas of Oral Drugs Launched 1983−2002[a]

therapy area	no. of drugs	indications listed in *Annu. Rep. Med. Chem.*
cardiovascular	79	antiarrythmic, anticholesterolemic, antihypertensive, antilipidemic, antithrombotic, calcium regulator, cardiostimulant, cerebral vasodilator, coronary vasodilator, diuretic, dyslipidemia, heart failure, hypertensive, hypocholeserolemic, hypolipidemic, platelet agg. inhib.
nervous system	74	Alzheimer's disease, analgesic, anti-Alzheimer, anticonvulsant, antidepressant, antiemetic, antiepileptic, antimigraine, antinauseant, antiobesity, anti-Parkinson, antiprolactin, antipsychotic, antispasmodic, antitussive, anxiolytic, cognition enhancer, convulsant, hypnotic, idiopathic hypersomnia, muscle relaxant, narcotic antagonist, neuroleptic, neuroprotective, nootropic
infection	64	antibiotic, antifungal, antimalarial, antiparasitic, antiviral
respiratory and inflammation	46	antiallergic, antiallergic/antiasthmatic, antiarthritic, antiasthma, antihistamine, antiinflammatory, antirheumatic, antisporiatic, bronchodilator, immunostimulant, immunosuppressant, organ rejection, antixerostomia, expectorant
gastrointestinal and metabolism	38	anabolic, antidiabetic, antiobesity, antiulcer, gastric antisecretory, antityrosinemia, gastroprokinetic, irritable bowel syndrome, ulcerative colitis, hyperprolactinemia, osteoporosis, Paget's disease
cancer	14	anticancer, antineoplastic
others	14	5α reductase inhib, abortifacient, antiglaucoma, antiprostatic hypertrophy, chelator, hypoammonuric, iron chelator, male sexual dysfunction, progestogen, urolithiasis, urologic, vit D prohormone

[a] Allocated from indications given in *Annual Reports in Medicinal Chemistry*.[15]

Table 4. Mean (Median) Physical Properties by Therapy Area of Oral Drugs, Launched 1983−2002[a]

	cardiovascular $n = 79$	nervous system $n = 74$	gastrointestinal and metabolism $n = 38$	infection $n = 64$	respiratory and inflammation $n = 46$	cancer $n = 14$	others $n = 14$
Mol Wt	389 (396)	310 (307)	378 (357)	456 (389)	396 (353)	313 (299)	309 (348)
cLogP	3.05 (3.00)	2.50 (2.55)	1.90 (2.28)	1.56 (0.94)	3.34 (2.90)	3.02 (3.01)	1.93 (2.22)
%PSA	19.8(18.6)	16.3 (14.3)	26.7(20.7)	24.6 (21.5)	20.5 (19.3)	20.8 (18.3)	22.9 (20.2)
OH + NH	1.46 (1)	1.50 (1)	2.71 (2)	2.41 (2)	1.37 (1)	1.00 (1)	1.64 (2)
O + N	6.73 (7)	4.32 (4)	6.84 (6)	8.78 (7)	6.17 (5)	4.5 (4.5)	4.29 (4)
HBA	3.77 (4)	2.12 (2)	4.34 (4)	5.28 (5)	4.24 (4)	2.86 (2.5)	2.64 (2)
RotB	8.23 (8)	4.70 (4.5)	7.63 (7)	6.83 (5)	5.52 (4.5)	5.00 (3.5)	4.57 (4.5)
Rings	2.84 (3)	2.85 (3)	2.32 (2.5)	3.45 (3)	3.02 (3)	2.36 (2)	2.36 (2.5)

[a] Mol Wt = molecular weight; cLogP = calculated 1-octanol/water partition coefficient (Daylight method); %PSA = calculated [(polar surface area)/(total surface area)] × 100; OH + NH = sum of OH + NH groups (H-bond donors); O + N = sum of O + N atoms; HBA = sum of H-bond acceptors; RotB = number of freely rotating bonds; Rings = number of rings in structure. Therapy areas were allocated according to Table 3. All physical property values are taken from ref 8.

Antiinfective drugs have the highest mean molecular weight and lowest mean lipophilicity, the highest O + N and H-bond acceptor count, and the largest number of rings. In addition, the distributions of molecular weight and lipophilicity among antiinfectives show different trends from the other drug classes as seen from the cumulative fraction plots (Figure 9). These observations with antiinfective agents are probably related to the need for activity in a nonhuman system, coupled with cell wall penetration requirements among antibiotic drugs. Nervous system drugs have significantly reduced molecular weight, polar properties (O + N, H-bond acceptors), and rotatable bonds relative to other classes. Most of these compounds are centrally acting, and their physical properties are consistent with the established importance of limited polar surface area and molecular size for blood−brain barrier penetration.[31] Apart from antiinfectives, the other main therapy area classes show similar distributions of lipophilicity values (Figure 9), reinforcing the overall importance of this property irrespective of therapy area or target mechanistic class. Respiratory and inflammation, and gastrointestinal and metabolism drugs have similar molecular weight profiles, but the distributions of cardiovascular and nervous system drugs show trends toward

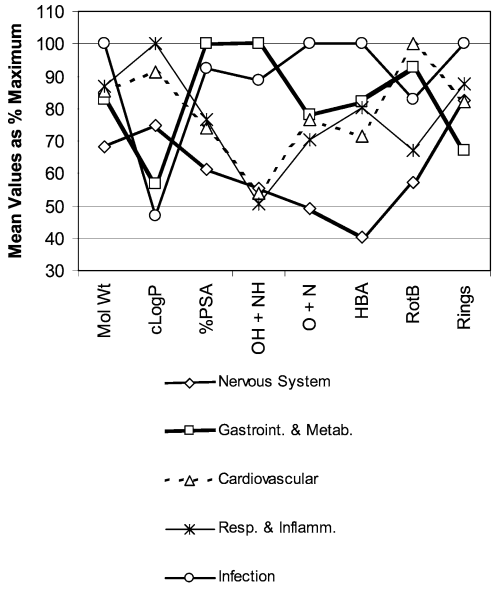

Figure 8. Relative physical property profiles of 1983−2002 oral drugs by therapeutic category.

respectively increased and decreased molecular weight relative to all other classes (Figure 9). The reduced

Table 5. Physical Properties Differing Significantly between the Five Principal Therapeutic Classes of 1983−2002 Oral Drugs[a]

	cardiovascular $n = 79$	nervous system $n = 74$	gastrointestinal and metabolism $n = 38$	infection $n = 64$
nervous system $n = 74$	**Mol Wt** %PSA **O + N** **HBA** **RotB**			
gastrointestinal and metabolism $n = 38$	cLogP *%PSA* *OH + NH* Rings	*Mol Wt* *%PSA* *OH + NH* **O + N** **HBA** **RotB** Rings		
infection $n = 64$	Mol Wt **cLogP** **%PSA** **OH + NH** **O + N** **HBA** *Rings*	**Mol Wt** cLogP **%PSA** **OH + NH** **O + N** **HBA** *RotB* *Rings*	Mol Wt *O + N* **Rings**	
respiratory and inflammation $n = 46$	**RotB**	*Mol Wt* %PSA *O + N* **HBA**	cLogP %PSA *OH + NH* RotB *Rings*	**cLogP** %PSA **OH + NH** *O + N* HBA

[a] Two-tailed p values were obtained from two-sample t-tests, assuming unequal variances. Properties in **bold** differ between the two therapeutic classes with $p < 0.001$, those in *italic* differ with $p = 0.001-0.01$, and other entries differ with $p = 0.01-0.05$.

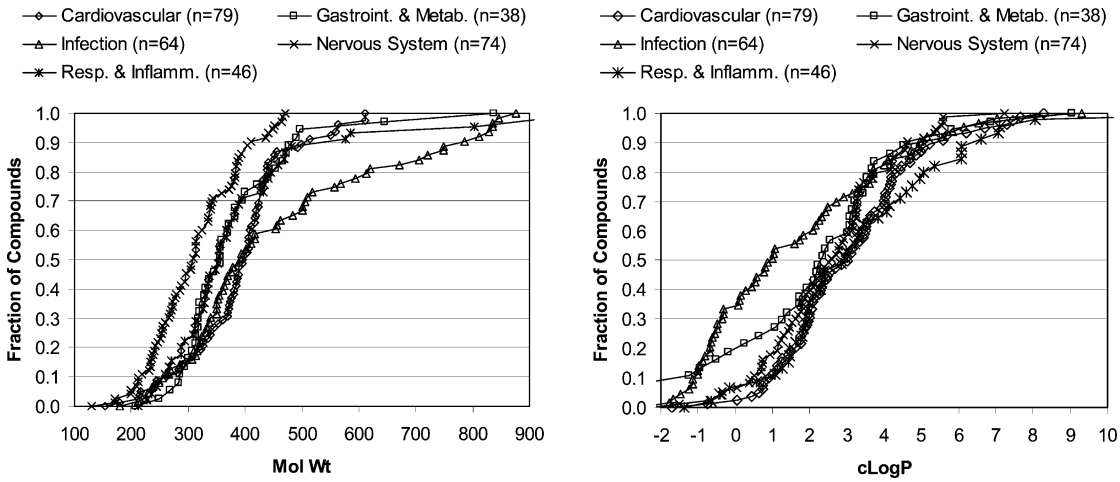

Figure 9. Cumulative fraction distributions of molecular weight (left) and lipophilicity (Daylight cLogP, right) in 1983−2002 oral drugs, according to therapeutic category.

variability of lipophilicity versus molecular weight distributions by therapy class is consistent with the overall profiles for older and newer drugs in Figures 3 and 4. Together these data support the view that lipophilicity appears to be a more stringent druglike property than molecular weight.

In each therapeutic area, except cardiovascular drugs, there has been no significant change in molecular weight with year of launch from 1983 to 2002 (Figure 10). Cardiovascular drugs have shown an increase in molecular weight over the same period (see Figure 10). A closer analysis of the cardiovascular drug group shows that 1983−2002 drugs have been dominated by only six different mechanisms: α and β adrenergic blockers, dihydropyridine-type calcium channel blockers, angiotensin converting enzyme (ACE) inhibitors, angiotensin II (AII) receptor antagonists, and hypolipidemic statin drugs. These clinically proven mechanisms account for 54 of the 79 drugs, and within each subclass there is clearly limited chemical diversity, since the drug molecules employ identical or very similar core pharmacophoric structures. There is a clear trend among these subclasses (expect α blockers, with only five drugs), showing an increased molecular weight with the year of first literature publication (see Figure 11), which helps to account for the overall increase in molecular weight with year of launch in cardiovascular drugs. This observation is not confined to cardiovascular drugs; among the 14 quinolone antibiotics introduced from 1983 to 2002 there is a 96 Da increase in molecular weight with time (data not shown). In this "derivative" approach to drug discovery, exploitation of the first breakthrough structure tends to result in increased size and structural complexity within each chemical class. These observations are consistent with other data showing that molecular weight increases during lead

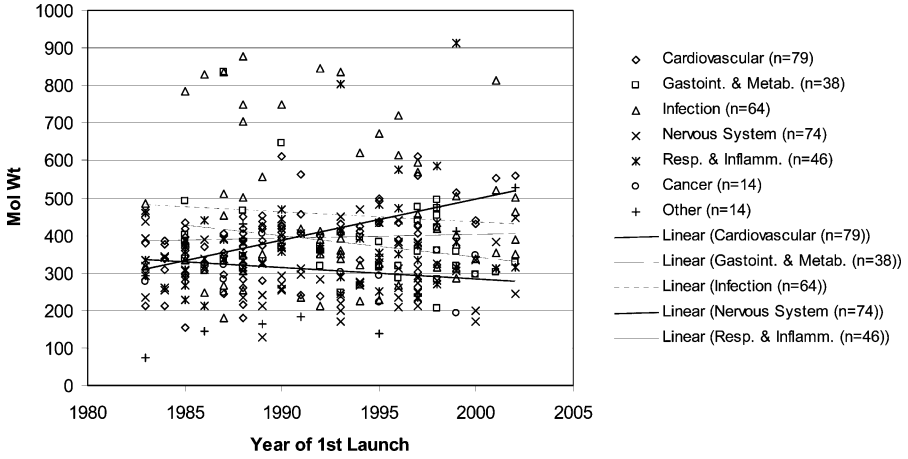

Figure 10. Plot of molecular weight versus year of first launch worldwide for 1983−2002 oral drugs, according to therapy area. Only cardiovascular drugs show a significant increase in molecular weight in this period ($n = 79$, $r^2 = 0.309$, $F = 34.4$, $p = 1.05 \times 10^{-7}$). All drugs show a nonsignificant upward trend of 1.2 Da per annum.

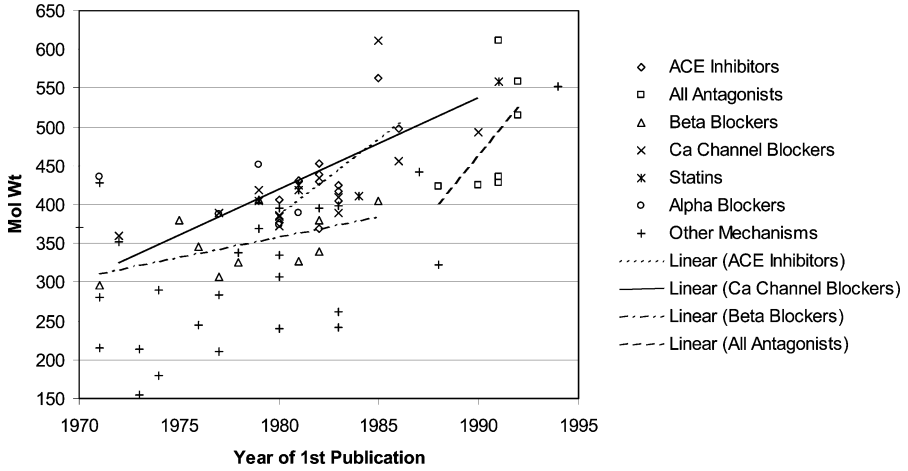

Figure 11. Plot of molecular weights of mechanistic classes of cardiovascular drugs introduced from 1983 to 2002 versus year of first literature publication.

optimization[26−28] and with the finding that a significant proportion of drugs approved in the year 2000 are derived from earlier drugs, a process that can lead to increases in physical properties in similar chemical classes.[32]

Conclusions

Mean values of lipophilicity, H-bond donors, and percent polar surface area are not changing between older (pre-1983) and newer (1983−2002) oral drugs. We propose that these are the most important oral druglike physical properties and should be carefully monitored and controlled in oral drug discovery programs. In contrast there has been a statistically significant increase in mean molecular weight, H-bond acceptors, rotatable bonds, and number of rings in newer drugs. Recent (1983−2002) oral drugs display a narrower range of allowable percent polar surface area, suggesting limitations may exist to increasing molecular size and structural complexity. The results suggest that the balance between polar and nonpolar drug properties is an important, unchanging feature of oral drug molecules. Although it is unclear how much the other physical properties (molecular weight, H-bond acceptors, rotatable bonds, and number of rings) can continue to be increased in future drug candidates, the data presented here on 1983−2002 drugs show that well-delineated ranges for these properties exist; therefore, their control should not be ignored. A comparison of physical properties by therapy area in drugs launched from 1983 to 2002 shows that antiinfective agents have a differing profile, being both larger and less lipophilic than other classes. Consistent with previous observations, nervous system drugs have reduced molecular weight compared with other therapy areas. The increase in molecular weight seen in 1983−2002 cardiovascular drugs is a consequence of capitalizing on several well-established and clinically proven mechanisms, employing limited chemical diversity. With the exception of antiinfective drugs, the need for an optimal drug lipophilicity profile appears to be independent of therapy area.

The data in this report show that the physical properties of oral drugs approved from 1983 to 2002 have increased by an average of 14% relative to pre-1983 drugs. A caveat of this retrospective analysis, reflecting drug discovery activities of around a decade ago at the latest, is that it does not necessarily provide a clear guide to future trends.[33] Approaches to address future drug "developability" have been emphasized strongly in the discovery phase over the past decade, including optimization of DMPK, physical form, and

safety properties in parallel with the desired biological effects. It is anticipated that these more stringent early selection procedures will result in reduced overall attrition rates in drug development. Since achieving these improved profiles will in general require more intensive lead optimization than in the past, with associated risks of further increases in molecular size and structural complexity, the control of physical properties within the druglike domain will remain a major challenge to drug design.

The structure of the chemical lead compound selected for optimization plays a pivotal part in determining the physical properties of the eventual candidate drug molecule. Consequently we suggest that lead generation efforts aimed at orally active drugs should be directed toward discovering a selection of diverse lower molecular weight leadlike[26,34] or fragmentlike[35] molecules as chemical starting points. These should ideally possess high binding efficiency for the desired target (expressed as binding energy per dalton[36]) and have the potential to achieve, following optimization, the best possible oral druglike profile. The selection of oral candidate drugs having optimal, rather than extreme, druglike physicochemical properties will provide additional confidence for successful drug development outcomes.

Acknowledgment. We thank our AstraZeneca R&D Charnwood colleagues Dave Cheshire, John Dixon, John Steele, Simon Teague and Mark Wenlock for numerous stimulating discussions, and we thank Michal Vieth (Lilly Research Laboratories) for providing Figure 1.

Supporting Information Available: Excel spreadsheet containing a table of oral drugs launched pre-1983 with physical properties, a table of oral drugs launched 1983−2002 with physical properties, year of launch, and therapeutic indication, and a table of results of pairwise comparisons of physical properties according to therapeutic area. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Frantz, S. 2003 Approvals: A Year Of Innovation and Upward Trends. *Nat. Rev. Drug Discovery* **2004**, *3*, 103−105.
(2) Source: 1999−2001 data from Centre for Medicines Research, International Institute for Regulatory Science.
(3) Reviews: (a) Muegge, I. Selection Criteria For Drug-Like Compounds. *Med. Res. Rev*. **2003**, *23*, 302−321. (b) Egan, W. J.; Walters, W. P.; Murcko, M. A. Guiding Molecules Towards Drug-Likeness. *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 540−549. (c) Lajiness, M. S.; Vieth, M.; Erickson, J. Molecular Properties That Influence Oral Drug-Like Behavior. *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 470−477.
(4) Takaoka, Y.; Endo, Y.; Yamanobe, S.; Kakinuma, H.; Okubo, T.; Shimazaki, Y.; Ota, T.; Sumiya, S.; Yoshikawa, K. Development of a Method for Evaluating Drug-Likeness and Ease of Synthesis Using a Data Set in Which Compounds Are Assigned Scores Based on Chemists' Intuition. *J. Chem. Inf. Comput. Sci*. **2003**, *43*, 1269−1275.
(5) 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.* **1997**, *23*, 3−25.
(6) Burton, P. S.; Goodwin, J. T.; Vidmar, T. J.; Amore, B. M. Predicting Drug Absorption: How Nature Made It a Difficult Problem. *J. Pharmacol. Exp. Ther*. **2002**, *303*, 889−895.
(7) Wenlock, M. C.; Austin, R. P.; Barton, P.; Davis, A. M.; Leeson, P. D. A Comparison of Physiochemical Property Profiles of Development and Marketed Oral Drugs. *J. Med. Chem*. **2003**, *46*, 1250−1256.
(8) Vieth, M.; Siegel, M. G.; Higgs, R. E.; Watson, I. A.; Robertson, D. H.; Savin, K. A.; Durst, G. L.; Hipskind, P. A. Characteristic Physical Properties and Structural Fragments of Marketed Oral Drugs. *J. Med. Chem*. **2004**, *47*, 224−232.
(9) Van De Waterbeemd, H.; Smith, D. A.; Beaumont, K.; Walker, D. K. Property-Based Design: Optimization of Drug Absorption and Pharmacokinetics. *J. Med. Chem.* **2001**, *44*, 1313−1333.
(10) Van De Waterbeemd, H.; Jones, B. C. Predicting Oral Absorption and Bioavailability. *Prog. Med. Chem*. **2003**, *4*, 1−59.
(11) (a) Blake, J. F. Examination of the Computed Molecular Properties of Compounds Selected for Clinical Development. *BioTechniques* **2003**, Suppl., 16−20. (b) 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) Kubinyi, H. Opinion: Drug Research: Myths, Hype and Reality. *Nat. Rev. Drug Discovery* **2003**, *2*, 665−668.
(13) 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.
(14) Sakaeda, T.; Okamura, N.; Nagata, S.; Yagami, T.; Horinouchi, M.; Okumura, K.; Yamashita, F.; Hashida, M. Molecular and Pharmacokinetic Properties of 222 Commercially Available Oral Drugs in Humans. *Biol. Pharm. Bull.* **2001**, *24*, 935−940.
(15) *Annual Reports in Medicinal Chemistry*; Doherty, A., Ed.; Academic Press: San Diego, CA, 2003; Vol. 38, pp 425−441 (lists 1983−2002 NCEs from the "To Market, To Market" chapters of previous volumes).
(16) (a) Cramer, R. D., III. BC(DEF) Parameters. 1. The Intrinsic Dimensionality of Intermolecular Interactions in the Liquid State. *J. Am. Chem. Soc*. **1980**, *102*, 1837−1849. (b) Van De Waterbeemd, H.; Smith, D. A.; Jones, B. C. Lipophilicity in PK Design: Methyl, Ethyl, Futile. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 273−286.
(17) Abraham, M. H.; Chadha, H. S.; Martins, F.; Mitchell, R. C.; Bradbury, M. W.; Gratton, J. A. Hydrogen Bonding Part 46: A Review of the Correlation and Prediction of Transport Properties by an LFER Method: Physicochemical Properties, Brain Penetration and Skin Permeability. *Pestic. Sci.* **1999**, *55*, 78−88.
(18) (a) Navia, M. A.; Chaturvedi, P. R. Design Principles for Orally Bioavailable Drugs. *Drug Discovery Today* **1996**, *1*, 179−189. (b) Chan, O. H.; Stewart, B. H. Physicochemical and Drug Delivery Considerations for Oral Drug Bioavailability. *Drug Discovery Today* **1996**, *1*, 461−473.
(19) Hansch, C.; Hoekman, D.; Leo, A.; Weininger, D.; Selassie, C. D. Chem-Bioinformatics: Comparative QSAR at the Interface between Chemistry and Biology. *Chem. Rev*. **2002**, *102,* 783−812. See Pomona Database of QSARs: Http://Clogp.Pomona.Edu/Medchem/Chem/Qsar-Db/Index.Html-Topic_Searching.
(20) Hansch, C.; Bjoerkroth, J. P.; Leo, A. Hydrophobicity and Central Nervous System Agents: On the Principle of Minimal Hydrophobicity in Drug Design. *J. Pharm. Sci.* **1987**, *76*, 663−87.
(21) Lewis, D. F. V.; Dickins, M. Baseline Lipophilicity Relationships in Human Cytochromes P450 Associated with Drug Metabolism. *Drug Metab. Rev.* **2003**, *35*, 1−18.
(22) Keseru, G. M. Prediction of hERG Potassium Channel Affinity by Traditional and Hologram QSAR Methods. *Bioorg. Med. Chem. Lett*. **2003**, *13*, 2773−2775.
(23) Goodwin, J. T.; Conradi, R. A.; Ho, N. F. H.; Burton, P. S. Physicochemical Determinants of Passive Membrane Permeability: Role of Solute Hydrogen-Bonding Potential and Volume. *J Med. Chem*. **2001**, *44*, 3721−3729.
(24) (a) Zhao, Y. H.; Abraham, M. H.; Hersey, A.; Luscombe, C. N. Quantitative Relationship between Rat Intestinal Absorption and Abraham Descriptors. *Eur. J. Med. Chem*. **2003**, *38*, 939−947. (b) Zhao, Y. H.; Abraham, M. H.; Le, J.; Hersey, A.; Luscombe, C. N.; Beck, G.; Sherborne, B.; Cooper, I. Rate-Limited Steps of Human Oral Absorption and QSAR Studies. *Pharm. Res.* **2002**, *19*, 1446−1457.
(25) Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. Hydrogen Bonding. 32. An Analysis of Water-Octanol and Water−Alkane Partitioning and the Delta Log P Parameter of Seiler. *J. Pharm. Sci*. **1994**, *83*, 1085−100.
(26) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. The Design of Leadlike Combinatorial Libraries. *Angew. Chem., Int. Ed.* **1999**, *38*, 3743−3748.
(27) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is There a Difference between Leads and Drugs? A Historical Perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308−1315.
(28) Hann, M. M.; Leach, A. R.; Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856−864.
(29) Rishton, G. M. Nonleadlikeness and Leadlikeness in Biochemical Screening. *Drug Discovery Today* **2003**, *8*, 86−96.
(30) Lipinski, C. A. Drug-Like Properties and the Causes of Poor Solubility and Poor Permeability. *J. Pharmacol. Toxicol. Methods* **2001**, *44*, 235−249. See also http://www.pharmalabauto.com/Solubility/Presentations/presentations.html for a number of related presentations by the author.
(31) (a) 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. (b) Hou, T. J.; Xu, X. J. ADME Evaluation in Drug Discovery. 3. Modeling Blood−Brain Barrier Partitioning Using Simple Molecular Descriptors. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 2137−2152 and references therein. (c) Doan, K. M. M.; Humphreys, J. E.; Webster, L. O.; Wring, S. A.; Shampine, L. J.; Serabjit-Singh, C. J.; Adkison, K. K.; Polli, J. W. Passive Permeability and P-Glycoprotein-Mediated Efflux Differentiate Central Nervous System (CNS) and Non-CNS Marketed Drugs. *J. Pharmacol. Exp. Ther.* **2002**, *303*, 1029−1037.

(32) Proudfoot, J. R. Drugs, Leads, and Drug-Likeness: An Analysis of Some Recently Launched Drugs. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1647−1650.

(33) From the list in ref 1, there are an additional 8 oral drugs approved for the first time in the U.S. in 2003 that are not included in this study (miglustat, epinastine, tadalafil, gefitinib, rosuvastatin, vardenafil, aprepitant, and atazanavir). The mean (median) physical property values for these 8 drugs show increases in physical properties relative to 1983-2002 oral drugs: Mol Wt 440 (464); cLogP 3.12 (2.85); O + N 7.75 (7) and OH + NH 2.38 (2). Their inclusion in the analysis does not alter the conclusions reached in this paper.

(34) Hann, M. M.; Oprea, T. I. Pursuing the Leadlikeness Concept in Pharmaceutical Research. *Curr. Opin. Chem. Biol.* **2004**, *8*, 255−263.

(35) (a) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-based Lead Discovery. *Nat. Rev. Drug Discovery* **2004**, *3*, 660−672. (b) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-Based Drug Discovery. *J. Med. Chem.* **2004**, *47*, 3463−3482.

(36) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The Maximal Affinity of Ligands. *Proc. Nat. Acad. Sci. U.S.A.* **1999**, *96*, 9997−10002.

JM049717D