

PHYSICOCHEMICAL PROPERTIES AND THE DISCOVERY OF ORALLY ACTIVE DRUGS: TECHNICAL AND PEOPLE ISSUES

CHRISTOPHER A. LIPINSKI

Pfizer Global Research and Development, Groton Laboratories,
Connecticut, 06340, USA

E-Mail: christopher_a_lipinski@groton.pfizer.com

Received: 13th June 2002 / Published: 15th May 2003

ABSTRACT

Poor aqueous solubility is the largest physicochemical problem hindering drug oral activity. Among combinatorial libraries, poor solubility is a frequently encountered problem but poor permeability is seldom a problem. The relative importance of poor solubility vs. poor permeability as a source of poor oral activity depends on the method of lead generation. Solubility or permeability problems are not purely a technical issue of assay design or computational prediction. People and organizational issues are extremely important. A computational ADME filter like the "rule of 5" (1) is most effective when used prior to the beginning of experimentation.

INTRODUCTION

Physicochemical property changes in recent drugs makes finding orally active drugs more difficult. Poor solubility will be viewed as the predominant problem if lead generation is heavily dependent on high throughput screening. Poor permeability will be viewed as the predominant problem if leads arise from structure based design. Adverse property changes can be managed through appropriate use of computational and experimental strategies. A computational filter for orally active drugs like the "rule of 5" is most effective when used prior to the beginning of experimentation because at this stage people issues are minimized. In my opinion, there is a hierarchy of properties that can be controlled by chemistry. Tight structure activity relationships (SAR) equate with good control. Properties important to oral activity like solubility and permeability do not show tight SAR and hence need early computational prediction and early experimental assays. Screening for poor aqueous solubility is important regardless of the type

of chemistry. It is important for both heterocyclic and peptido-mimetic compounds. Medium to high throughput solubility assays, for example turbidimetric solubility assays, are only useful in early discovery. Traditional thermodynamic solubility assays are most appropriate to the discovery development interface when the crystalline state of drugs is well characterized. In contrast to the screening for poor aqueous solubility, the value of screening for poor permeability depends much more on the chemistry chemo-type. Experimental permeability screening is most valuable for conformationally flexible compounds particularly those containing multiple charged groups. By contrast permeability screening for heterocyclic compounds particularly those containing few rotatable bonds may not be very useful unless the permeability problem is related to a biological transporter. Heterocyclic compounds containing few rotatable bonds are the frequent products of combinatorial chemistry and computational predictors for permeability suggest that few compounds in combinatorial libraries will exhibit a permeability problem.

METHOD

There is a systematic method to understand the causes and potential solutions to problems of poor physicochemical properties that are associated with poor oral absorption. This method involves a historical and database analysis on how physicochemical properties have changed with time from the era where problems with poor oral absorption were not so pronounced. This chemo informatic database method is very analogous to the rationale often given for studying history. In the affairs of man it is necessary to understand the past (history) to avoid in the future repeating the errors of the past.

Two technologies in lead discovery have to a considerable extent dominated the scenario of drug lead generation. These technologies are high throughput screening (HTS) and combinatorial chemistry. It is very easy to track the onset of these technologies by performing a simple citation analysis. I searched SciFinder 2001[®] software from Chemical Abstracts using the text string “high throughput screening” and the text string “combinatorial chemistry”. Both searches gave essentially an identical profile with a rapid increase in literature citations starting in about 1995. The similarity in the citation profiles is very reasonable. HTS is the rapid screening of large numbers of compounds in a biological assay. The biological screening process requires large numbers of chemistry compounds to be assayed. Combinatorial chemistry, the automated generally robotic synthesis of large numbers of chemistry compounds

provides the material to be screened in the HTS assays. It is common today to find statements in magazine articles similar to the following “HTS and combinatorial chemistry have not lived up to their promise”. These statements are partly true but misleading because they fail to differentiate between the early and later stage of combinatorial chemistry and whether the problem is in the HTS process or the combinatorial chemistry screening file. In my opinion there is not a problem with HTS. The problem lies in the fact that the first fifty percent of the history of combinatorial chemistry was badly flawed from an oral drug delivery perspective. The valid technology of HTS could not easily yield drug-like (orally active) drug matter if the combinatorial chemistry compound starting points were badly flawed.

There are two factors responsible for the production of badly flawed combinatorial compounds up to about the 1997-8 time period. The earliest factor in a time sense leading to the production of badly flawed combinatorial compounds was the actual method of robotic chemistry synthesis. A new technology tends to adapt pre-existing technology. In the case of combinatorial chemistry the pre-existing technology was the Merrifield solid phase synthesis of peptides. This automated method of peptide synthesis was in place before the advent of combinatorial chemistry and automated synthesis equipment was commercially available. Peptide scaffolds are capable of presenting interesting chemistry functionality in various regions of space and so the earliest combinatorial libraries were constructed using peptide scaffolds. Initially the work focused on naturally occurring α -amino acids and later with non-natural amino acids. Early workers were fascinated with the possibility of discovering compounds with potent *in vitro* activity. This focus was completely understandable given the difficulty of discovering a drug lead with potent *in vitro* activity in the decade of the 1980's. Peptide scaffold based combinatorial libraries did indeed generate potent *in vitro* active compounds in the new HTS screens but it took a number of years to realize that these initial HTS hits were very difficult to convert into orally active compounds. Naturally occurring α -amino acid bonds are metabolically unstable so these early peptide based libraries had little or no *in vivo* activity. Another problem that was initially not appreciated is that a compound with more than just a few amide bonds can be quite impermeable through the gastrointestinal wall. Hence many of these early peptide scaffold based combinatorial libraries were very poorly absorbed by the oral dosing route.

The second factor responsible for the production of badly flawed combinatorial compounds up to about the 1997-8 time period can be traced to the inappropriate implementation of the concept

of maximum chemical diversity. In the concept of maximum chemical diversity one tries to synthesize compounds with interesting chemical functionality displayed in as many directions as possible in three dimensional space. The idea is to display chemistry functionality likely to be involved in target recognition in as many areas of chemistry space as possible. The greater the coverage of chemistry space with appropriate chemistry functionality the greater the likelihood of detecting activity in an HTS assay. Initially workers did not know how much chemistry functionality was necessary. It seemed likely that more was better. For example building a compound from four fragments gave a greater display of functionality than building a compound from three fragments. Also the theoretical number of combinatorial compounds that could be produced from four fragments was much larger than from three fragments. This was attractive because of the logic that screening greater numbers of compounds increased the probability of finding an active hit in an HTS assay. Hence many combinatorial libraries (collections of compounds) were synthesized from four fragments. Again early workers were fascinated with the possibility of discovering compounds with potent *in vitro* activity and HTS screening of these early tetramer libraries did indeed result in HTS hits with potent *in vitro* activity. It took a period of time before researchers discovered that these potent *in vitro* tetramer library hits were not producing orally active compounds on subsequent medicinal chemistry optimization. The problem was that the average tetramer combinatorial compound is very large with a molecular weight perhaps in the 650 Dalton range. Compounds in this molecular weight range tend to be both very impermeable through the gastrointestinal wall and very insoluble. Hence the phenomenon of potent *in vitro* activity but very poor or no *in-vivo* activity was observed.

Several more minor factors exacerbated the reliance of most pharmaceutical companies on these early flawed combinatorial libraries. The Pfizer Groton Connecticut laboratories began HTS in the late 1980's before the advent of combinatorial chemistry. Realizing the need for massive numbers of compounds for HTS screening Pfizer began a massive campaign to purchase available compounds from academic laboratories. This effort was well funded and very successful and largely completed by 1994. As a result, purchase of academic compounds was not a viable option by the time other pharmaceutical companies realized the need for acquisition of large numbers of compounds for HTS. Pfizer had quite literally cleaned out the world wide academic supply. A second factor exacerbating the reliance of most pharmaceutical companies on these early flawed combinatorial libraries was the unreliability of the only remaining

alternative compound source to combinatorial chemistry that existed in the early 1990's. In 1991 the Soviet Union ceased to exist and quite rapidly large numbers of Soviet block chemists became unemployed and had to feed themselves and their families. A large synthetic capacity existed in the former Soviet Union and a demand for compound supply certainly existed in the largely western pharmaceutical companies. This supply and demand should in theory have resulted in a good match between supplier and customer. Unfortunately the quality control of these early Soviet block compounds was very poor. For example, in our Pfizer experience with these compounds we quite literally encountered the same compound sold to us with five different chemical structures. In our case this problem was not resolved until the compounds were delivered with appropriate spectral proof of identity. I believe our experience was probably shared by other companies. The result was that HTS screening results from these early Soviet block compounds was quite unreliable. In recent times the situation has completely changed. High quality compounds both combinatorial and non combinatorial accompanied by spectral documentation are now available from various vendors from the former Soviet block countries.

Aqueous solubility and permeability data must be provided to chemistry as early as possible to avoid oral absorption problems.

The 3D graph (Fig. 1) illustrates the three parameters under chemistry control that determine whether a compounds physicochemical profile is compatible with oral activity. The chemist has to synthesize a compound to achieve the appropriate combination of potency, solubility and permeability to move the compound into the region of space occupied by an orally active compound (above the solid surface). The points below the surface represent possible starting points in a lead optimization process. Usually the starting point is inferior in all three properties. Very frequently if only potency is improved it may be impossible to achieve oral activity (even with high potency) if solubility and permeability are very poor. The optimization of potency at the expense of poor solubility and / or poor permeability is a common occurrence Medicinal chemistry *in vitro* potency improvement usually does not improve a solubility or permeability problem in the lead starting point. In fact the usual pattern is that *in vitro* activity optimization results in an increase in both molecular weight and lipophilicity. Increases in these properties tend to correlate with increased poor aqueous solubility. In theory, extremely high potency will solve a permeability or solubility problem. In practice, it is quite difficult to get orally active drugs at doses below 0.1 mg/kg.

Aqueous Solubility and Permeability Data Must be Provided to Chemistry as Early as Possible to Avoid Oral Absorption Problems

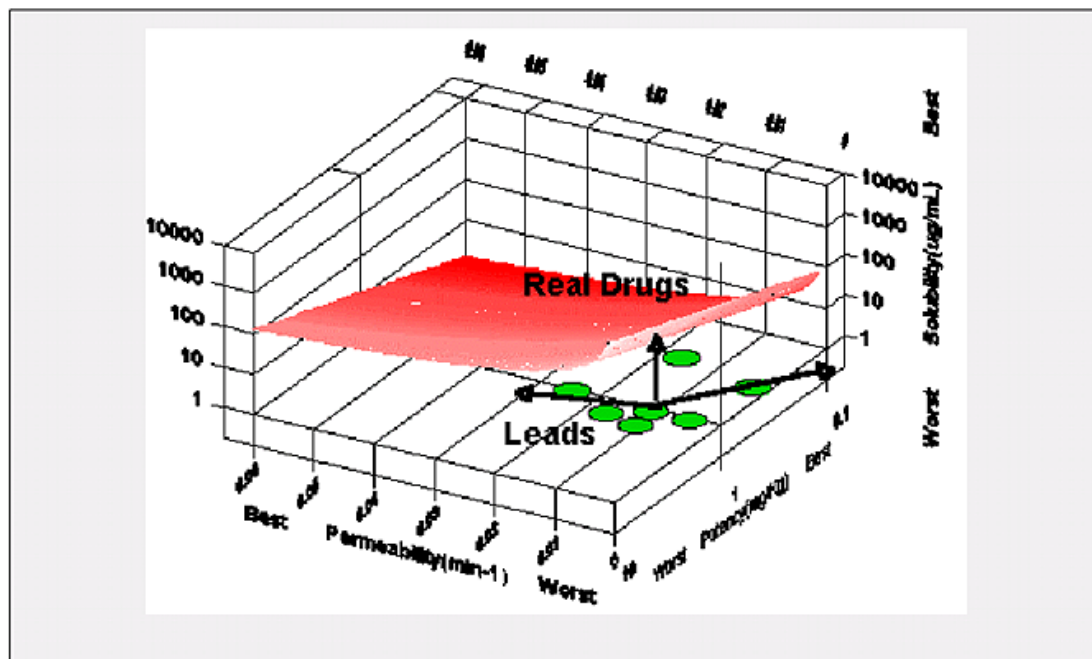


Figure 1.

The reason is that at very low doses a variety of metabolic processes can easily destroy the drug. At higher drug doses, these metabolic processes are saturated and less important. In my opinion, it is often easier to solve solubility problems than to solve problems in passive membrane permeability since the range in drug-like solubility is much greater than for permeability. For example, the FDA's proposed bio-equivalence classification system (BCS) classifies drugs into 4 classes depending on whether drugs have high or low permeability and high or low solubility. In the BCS, the range for permeability covers considerably less than three orders of magnitude while that for solubility covers a full six orders of magnitude. The best way to solve a permeability or solubility problem is with chemistry. The key to avoiding this problem is to provide the chemist with information on solubility and permeability at the same time as the potency information is received.

The General Pharmaceuticals Laboratory in our development organization profiles all newly nominated clinical candidates. As part of the evaluation, a minimum absorbable dose (MAD) is calculated for oral dosage forms based on the expected clinical potency, the solubility and the permeability. This calculation serves to confirm that either the physicochemical properties of

the candidate are easily within the acceptable range or that the properties lie within a difficult range that will require more than the average pharmaceuticals manning to solve any difficulties.

We have adapted this calculation to create a simple bar graph (Fig. 2) that we distribute to our medicinal chemists. It answers the question of “how much solubility does the chemist need?” Presented in bar graph format the information is very readily understood by our chemists. Presented to our chemists in the original published equation format its impact on our chemists was poor. Bar graph (2) illustrates a people issue. It is intended for presentation to our medicinal chemists and uses information from a paper published by a Pfizer pharmaceutical sciences researcher on a minimum absorbable dose (MAD) calculation. It is very important to present information in a format readily grasped by the intended audience. Pharmaceutical scientists are very comfortable with information presented in an equation format. Synthetic organic chemists are uncomfortable with mathematic equations.

There is a simple reason for this. American Ph.D. granting programs require four semesters of calculus to obtain a Ph.D. in chemistry. However calculus is not needed to be a competent synthetic chemist. All that is really needed is the mathematical skill set that typically comes from a quality high school education. Synthetic chemists tend to forget those math skills that are not needed in their profession.

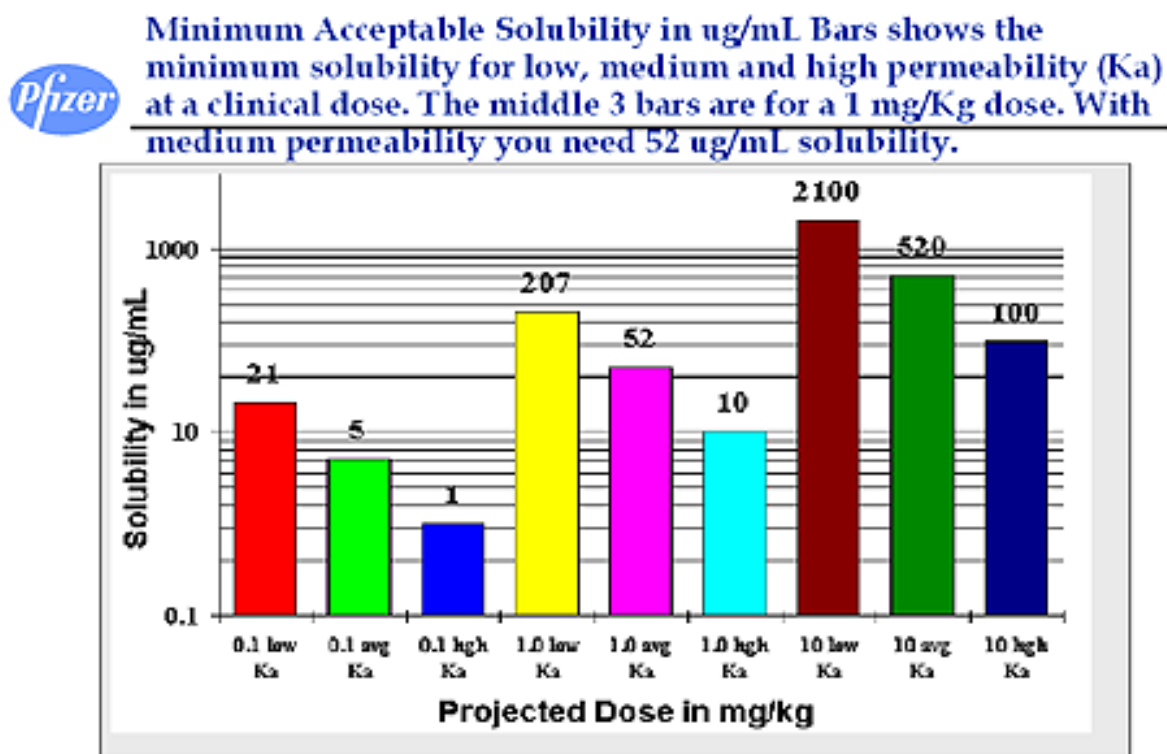


Figure 2.

By way of contrast synthetic chemists have very finely tuned pattern recognition skills with the ability to read a tremendous amount of information from a chemical structure depiction. This pattern recognition skill carries over to a graphical representation like a bar graph.

Minimum Acceptable Solubility for a drug can be calculated using an equation that takes into account the drug dose (potency), the solubility, the anticipated permeability and the intestinal fluid volume (assumed to be a constant). The usual solubility concentration units are $\mu\text{g/ml}$. For a molecular weight of 500 Daltons 5 $\mu\text{g/ml}$ translates to a molar concentration of 10 μM . The acceptable solubility ranges are displayed in bar graph (2). Each set of three bars shows the minimum solubility for compounds with low, medium and high permeability (K_a) at an anticipated clinical dose. The middle set of three bars is for a 1 mg/kg dose. With medium permeability you need 52 $\mu\text{g/ml}$ solubility. The three middle bars describe the most common clinical potency that we encounter; namely that of 1 mg/kg. If the permeability is in the middle range as for the average heterocyclic (the purple bar) then a thermodynamic solubility of about 50 $\mu\text{g/ml}$ at pH 6.5 or 7 is required. If the permeability is low (as in a typical peptido-mimetic) the solubility should be about 200 $\mu\text{g/ml}$.

Leads at Pfizer and in the drug industry in general, now trend toward higher molecular weight and lipophilicity. Bar graph (3) shows the trend in molecular weight for compounds synthesized in our medicinal chemistry labs (shown in red) and compounds purchased from external commercial sources (shown in blue). In our Pfizer Groton laboratories we began HTS screening in 1989, and increased HTS through 1992. The percentage of compounds with a molecular weight over 500 (which we believe is undesirable) tracks exactly with the increased HTS screening. More and more of our leads were from HTS, these had poorer physicochemical profiles and when our medicinal chemists followed up these leads they made compounds with profiles like those of the leads or sometimes even worse than those of the leads. The trends in compounds made in our medicinal chemistry labs are not aberrant; they are completely logical (and predictable) in terms of medicinal chemistry principles and the information available to the chemists. For example, introducing a lipophilic moiety (e.g. a methyl) so as to fit into a receptor is one of the best ways to improve *in vitro* potency. This same change however, also increases lipophilicity. Compounds purchased from commercial sources (in blue) were intended for random HTS screening and show no upwards trend in high MWT. A bar graph with high lipophilicity instead of high molecular weight would look very similar.

Computationally comparing libraries allows one to deduce the differences between real drugs and those medicinal chemistry compounds which do not really possess drug-like properties. One can use the presence of an International Non-Proprietary Name (INN name) or a United States Adopted Name (USAN name) or marketed status as a marker for a compound with “drug-like” properties. Inn names and USAN names are assigned when a compound enters phase two clinical efficacy studies. Entry into clinical phase II efficacy study is a marker for drug-like properties.

Compounds that fail to survive the phase I human toleration studies or the pre-clinical stage do not receive an INN or USAN name. The compounds with severe oral absorption, toxicity and metabolism issues have been filtered out in a compound achieving phase II status. A compound entering into phase II is a real drug in the sense that except possibly for efficacy it has all the attributes of a real drug. Historically, of those drugs reaching phase II, ninety percent have been intended for oral administration.



Leads at Pfizer and in the drug industry in general, now trend toward higher MWT and lipophilicity

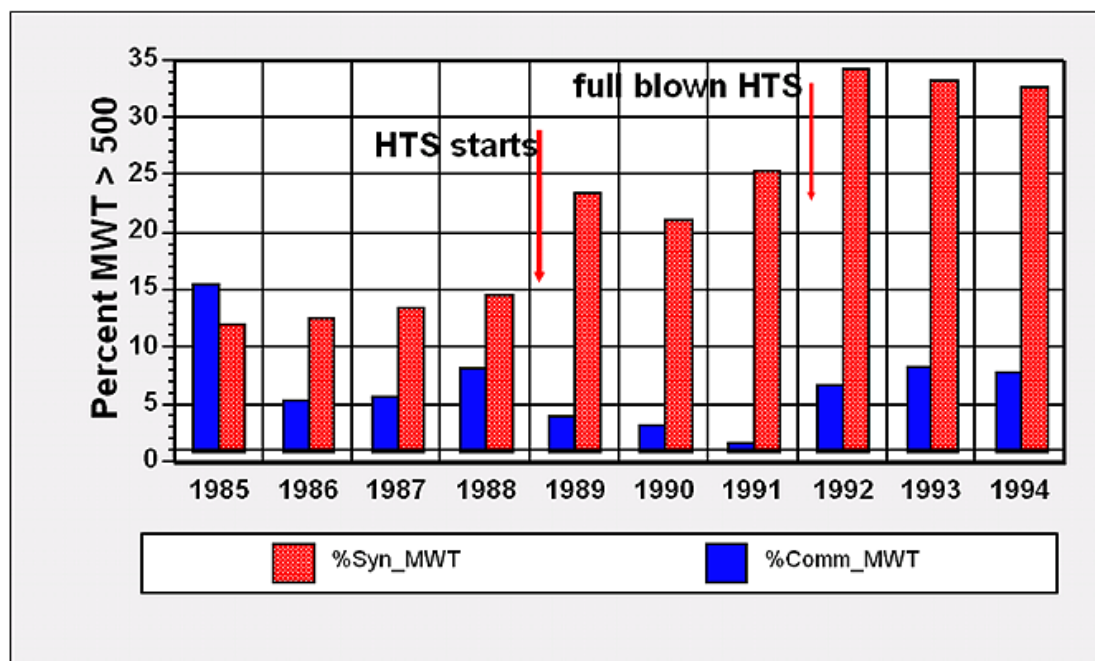


Figure 3.

So the presence of an INN or USAN name reliably identifies a set of orally active drugs with “real” drug like properties.

Drug-like as opposed to non drug-like physicochemical characteristics can now be defined by comparing drug-like with non drug-like data sets. The drug-like data set is a set of 7483 drugs which encompasses drugs with an INN/USAN name as well as drugs that were actually approved for marketing in at least one country. The library of 7483 INN/USAN and marketed drugs that was our benchmark is a significant fraction of all drugs that have reached phase II status. For example, there were about 9,500 USAN drugs listed in the most recent publication of the US Pharmacopeia.

The non drug-like data set is a set of 2679 drugs with the character that they represent a much earlier stage of drug discovery before any significant filtering for drug-like properties has occurred. I obtained this data set from the Derwent World Drug Index using the following procedure. I looked for drugs where the mechanism field contained the text “trial preparations”. This procedure identified drugs intended for a medicinal therapeutic purpose. I excluded any drug with a Chemical Abstracts Service (CAS) registry number. This effectively made sure that the drug had only recently been abstracted into the Derwent World Drug Index (WDI) because I knew from experience that it typically took about two years for the CAS registry number to be included in the WDI. I also double checked that no compound in the non drug-like data set had an INN/USAN name. The compounds in the non drug-like data set were all abstracted in 1997, 1998 and 1999. This data set will of course contain some real drug-like compounds but it will also contain many more compounds that are only ligands for a biological target and that lack some or all of the attributes required for an orally active drug. This non drug-like data set represents the type of early discovery stage compound that one is likely to encounter prior to any filtering operation. Compounds similar to this data set are likely to be encountered in preliminary reports of biological activity at scientific meetings and to appear in the primary medicinal chemistry literature.

I have compared the distribution of the physicochemical properties for the drug-like compounds and the non drug-like compounds in figure 4. The reader can also think of these data sets as corresponding to the newer (non drug-like) and older (drug-like) compounds.

A convenient method of comparing the distribution of a property across non equally sized data sets is to compare Kaplan-Meier type survival curves. This graph shows the distribution of molecular (formula) weights of four classes of compounds. Shown in blue are drugs with International Non Proprietary Names (INN) and United States Adopted Name (USAN) name. These are the drug-like compounds that have survived phase I with sufficient oral

bioavailability and acceptable pharmacokinetic and pharmacodynamic parameters to reach phase II. Shown in green are New Chemical Entities (NCE). These are the drugs that actually reached market and are the compounds that are summarized in the “To Market - To Market” chapter in the back of the issues of “Annual Reports in Medicinal Chemistry”.

By definition these are certainly real drugs. Shown in yellow are compounds appearing in the Derwent World Drug Index. This includes a very wide range of compounds. All have some sort of biological activity. Shown in red are the new drugs (the non drug-like) data set of drugs that were abstracted by Derwent in 1997, 1998 and 1999 from the medicinal chemistry journal and conference literature. The MWT corresponding to the 90th percentile and a decreasing probability of oral activity is marked by a horizontal line.

Newer drugs (red) are larger

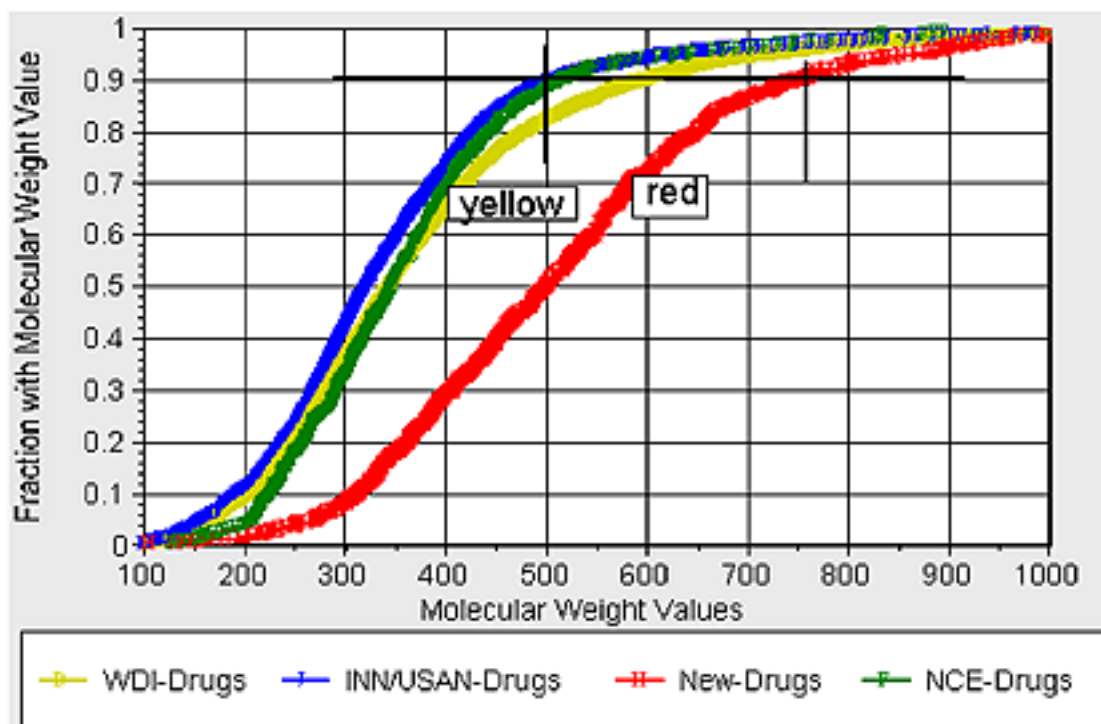


Figure 4.

When a curve is shifted towards the right it means that globally that data set has a higher distribution of the parameter. The red curve is distinctly shifted toward the right relative to all the other curves. This means that the new drugs overall have higher molecular weights than the real drug-like compounds. Newer drugs (non drug-like compounds) are larger in size than

traditional, older real drugs. Figure 5 shows a set of Kaplan-Meier like curves for four physicochemical parameters in the same INN / USAN data set.

The idea is that the distribution of parameters for INN / USAN drugs can be used to define a property range where oral activity is increasingly difficult due to poor absorption or poor permeability. The distribution of four parameters for 7483 INN/USAN drugs define the ninety percent limits corresponding to properties unfavorable for oral drug absorption. The four properties were chosen based on extensive literature precedent. Too high a molecular weight was previously known to be linked to poor solubility and permeability. It was previously known that typically for a particular drug series there was an optimum lipophilicity for biological activity.

Percent Distribution of Drug Like Parameters

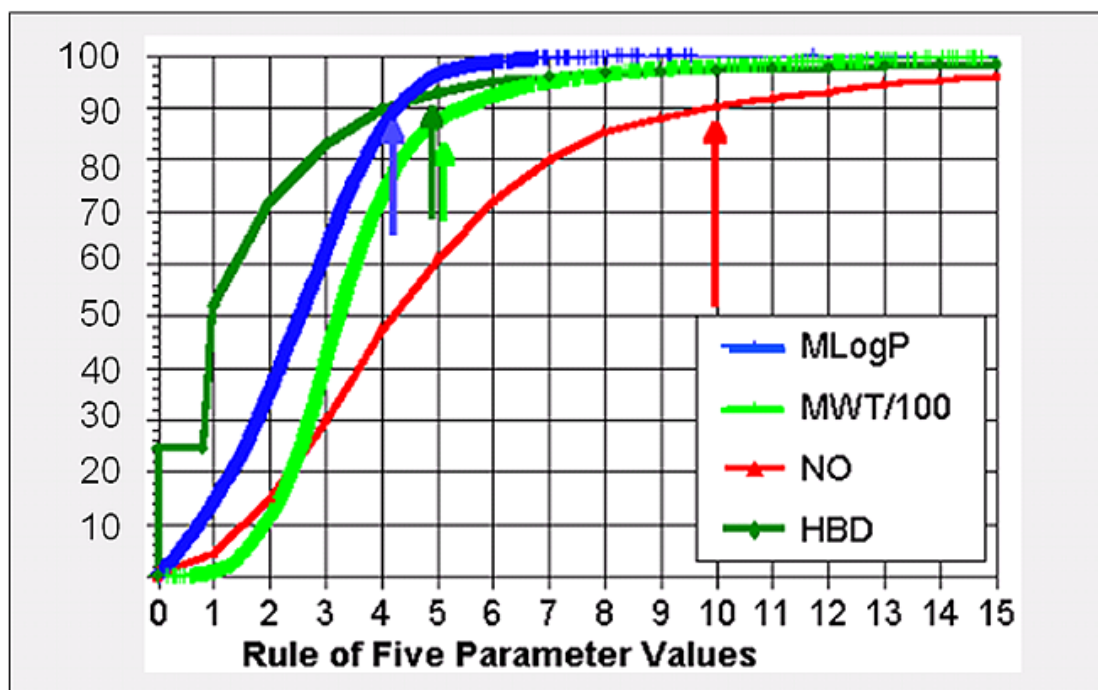


Figure 5.

Too little or too much lipophilicity was detrimental to biological activity. From work on peptides and peptide-mimetic compounds it was known that too many hydrogen bonding interactions between drug and water were detrimental to the ability of the drug to cross (permeate) the gastrointestinal wall. Permeation of the gastrointestinal wall is an absolute requirement for oral activity in a drug

All the curves exhibit a leveling as parameters reach unfavorable values for oral activity. The 90th percentile of each parameter is shown by the arrows. The colored lines show the distribution of: MLogP - lipophilicity (in blue) as measured by the Moriguchi Log P algorithm; MWT/100 - molecular weight (in light green), divided by 100 for plotting; OH+NH - the sum of OH plus NH (in red) as an index of H-bond donors; O+N - the sum of oxygen plus nitrogen (dark green) as an index of H-bond acceptors. There is a very clear similarity in the patterns of all four curves. For each parameter most of the values lie in the region below the ninety percent asymptote. It appears as if for real drugs that there are limiting values for all four of these parameters. The vast majority (ninety percent) of these real drug-like compounds do not exceed a particular parameter value.

RESULTS

This analysis led to a simple mnemonic which I called the “Rule of 5” because the parameter cutoff values all contained 5’s. Numerically there actually are only four rules.

The “Rule of 5” states: Poor absorption or permeation are more likely when there are:

- More than 5 H-bond donors
- The MWT is over 500
- The CLog P is over 5 (or MLOGP is over 4.15)
- The sum of N’s and O’s is over 10
- Substrates for transporters and natural products are exceptions

Although this rule is very simple, it works remarkably well provided you understand its limitations. First, it only works because the physical property profile of medicinal compounds being currently made is quite far outside that of marketed drugs. Secondly, it doesn't work for compounds that are of natural product origin or have structural features originally derived from natural products, for example antibiotics, antifungals. The likely reason is the important roles of transporters in these classes. It also doesn't work well for certain therapeutic areas where many drugs are substrates for biological transporters. Anti-infective agents are a specific example of a therapeutic class where the “rule of 5” does not work well. Many anti-infective agents, e.g. the orally active cephalosporins are orally active because they are substrates of the PEPT-1 biological transporter. The affinity for the biological absorptive transporter allows the drug to bypass the physicochemical “rule of 5” limits for gastrointestinal wall permeability. Pfizer uses the “Rule of 5” in a variety of ways. For example it is used as an on-line alert at

compound registration. It is used as a filter for high throughput screening (HTS) libraries. We do not screen libraries (collections) of compounds with significant non-compliance to the “rule of 5”. We use the “rule of 5” as a filter for purchased compounds. We use it as a criterion for focused library synthesis and we use it as a guideline for quality clinical candidates. The latter use is not unique to Pfizer. In fact more and more there is recognition in the literature that the quality of the starting point in a chemistry optimization process is a good index of the final quality of the clinical candidate. We are now seeing analyses where researchers are tracking the relationship between the structure of a marketed drug to the structure of the starting point leading to the drug. The tight relationship between the starting point and the final drug is remarkable. A good starting point is likely to lead to a good drug. Conversely it is very difficult (but not impossible) to convert a poor starting point into a quality drug clinical candidate.

Considerable information relating to possible causes of poor solubility and poor permeability results from looking at how the properties of a drug company clinical candidates have changed with time (2). I am going to compare how important properties have changed with time for two very different drug organizations by comparing the properties of clinical candidates from the Pfizer Groton CT labs and the worldwide Merck organization.

Both organizations have been very successful in discovering drugs but they have done it in very different ways. For example one can plot the molecular weight (essentially a measure of size) for each early stage clinical candidate from Pfizer’s Groton labs and fit the best straight line through the points. One sees lots of scatter but the trend is clearly up. Over the years Pfizer Groton clinical candidates have gradually become larger. One can also discern the industry wide trend towards higher molecular weight in clinical candidates from Merck by analyzing the molecular weight trend with time for Merck advanced candidates (identified by MK numbers). Merck MK-numbers are issued in non sequential order and not all Merck MK compounds in the literature are candidates. For example, important biological standards may be assigned an MK-number. For this reason, the time scale for the analysis is the date of the earliest Merck patent corresponding to the MK-number candidate.

One can examine the trend of MWT as a function of time for each Merck candidate and fit the best trend line. Although there is considerable scatter there is clearly an upward trend in molecular weight with time. So just like Pfizer, Groton, Merck’s clinical candidates have also gotten bigger with time. In a similar exercise one can plot the lipophilicity trend with time for Pfizer Groton clinical candidates. There is an upward trend with time for Pfizer Groton clinical

candidates to become more lipophilic. It appears as if they are pushing right up to about a limit of 4-5 in logP. They don't go much higher because it really gets hard to get an orally active drug when you exceed a value of 4-5 (the specific value of Log P varies a bit with the method of calculation). This just does not happen with clinical candidates from Merck. With time they absolutely do not become more lipophilic. So there has to be something very different about how Pfizer and Merck discover drugs. Not better or worse, just different.

Hydrogen bond acceptors are atoms in a drug that can accept an interaction with water through a hydrogen bond. Too few hydrogen bonds in many cases is not a good thing and too many hydrogen bonds is also not a good thing. More than about ten hydrogen bonds in a drug is not good because the drug will have difficulty getting through the wall of the intestine into the blood. A drug given by mouth (an orally active drug) has to get from the inside of the intestinal tract through the intestinal wall to reach the blood stream. Certain kinds of atoms like Nitrogen (N) and Oxygen (O) in a drug accept these hydrogen bonds. So if you just count up the number of N's and O's in the drug molecular formula you get a simple (but still quite useful) measure of this hydrogen bond accepting property. There is a trend with time towards increasing number of hydrogen bond acceptors among Merck candidates. This trend is what one might expect given the strong focus in structure based drug design in recent years and on a type of chemistry called peptidomimetic like structures. This is the kind of drug discovery that Merck is famous for and very good at. A similar analysis for Pfizer Groton candidates would absolutely not show this upwards trend in hydrogen bond acceptors. Pfizer Groton does a lot of HTS whereas over the time period of this analysis Merck focused more heavily on all the various approaches to rational drug design other than HTS). One approach is not better than the other, just different. But the differences in approaches show up over time. So there must be something about HTS as opposed to non HTS drug discovery that leads to differences in trends towards increased hydrogen bonding functionality over time. So what do these differences between Pfizer, Groton and Merck and the differences in clinical candidate trends with time mean? Well with Merck the trend is towards larger size and more hydrogen bond interactions between the drug and water. Taken too far this translates to a problem in getting through the gastrointestinal tract wall. So an organization like Merck tends to worry about this property of poor permeability (problems getting through the gut wall). With Pfizer in Groton the trend is towards larger size and greater lipophilicity. Taken too far this translates to a problem in dissolving in the water

inside the gastro intestinal tract. This is a problem of poor solubility and a drug has to be soluble to be orally active.

There is no free ride in drug research. Every discovery approach has a downside. Poor solubility and poor permeability are both bad. But they are not equally bad. It is much better to have poor solubility than poor permeability. The reason is that currently there are pharmaceutical sciences fixes for poor solubility. One would like to avoid them for all kinds of reasons but they do exist. By contrast, there is no pharmaceutical sciences formulation fix for poor permeability (except changing the chemistry structure as in a pro-drug) and there likely will not be for at least the next five years.

What are the reasons for the different physicochemical profiles in structure based as opposed to HTS based discovery approaches? In structure based approaches one is typically working on enzyme inhibitors or peptido-mimetics. Potency enhancement usually involves probing for at least three binding sites, e.g. in the P1, P1', P2 pocket. The binding pocket is often elongated. These considerations tend to lead towards larger size. Hydrogen bonding count tends to go up because one is often trying to satisfy multiple receptor hydrogen bonding interactions. Often the natural ligand is a peptide. There is not much selection pressure for log P to increase because a lot is known about the target. Lipophilicity does not play a role in discovering the lead series as it does in the HTS based discovery approach. Large size and increased H-bonding translates to a poorer permeability profile. HTS based approaches tend to bias towards larger size and higher lipophilicity because these are the parameters whose increase is globally associated in a medicinal chemistry sense with an improvement of *in vitro* activity. Larger size and higher lipophilicity translate to poorer aqueous solubility. Fortunately for HTS based approaches this bias can be corrected by appropriate filtering based on compound physicochemical properties.

Combinatorial libraries show a distinctive pattern with regards to permeability and solubility. Specifically, poor permeability is seldom encountered as a problem in a combinatorial library. As a result permeability profiles are not very dependent on chemistry synthesis protocol. Almost any protocol will result in compounds predicted to have an acceptable permeability profile. The reason relates to chemistry. It is actually quite difficult to construct a combinatorial library with permeability problems. It is difficult to make libraries with many hydrogen bond donors and acceptors in a combinatorial manner. To illustrate this point I analyzed a set of 47,680 combinatorial compounds from the same commercial source made according to 30 different synthesis protocols. There was little variation in average polar surface area (PSA)

across the protocols and all but one of the thirty protocols gave an average PSA of less than 140 square Angstroms. A PSA of less than 140 square Angstroms suggests that passive trans membrane permeability (the most common drug permeability mechanism) will be quite acceptable.

Solubility profiles in contrast to permeability profiles can be very dependent on the chemistry synthesis protocol. The reason is that poor aqueous solubility is the major physicochemical problem found in combinatorial libraries. I calculated average aqueous solubility in $\mu\text{g/ml}$ for the same data set of 47,680 combinatorial compounds from the same commercial source made according to 30 different synthesis protocols. The solubility program I used was a Pfizer developed model based on experimental data in our discovery turbidimetric solubility assay. The experimental assay outputs solubility in aqueous pH 7.0 phosphate buffer in the range < 5 to > 65 $\mu\text{g/ml}$ using a turbidimetric end point. The computational model based on experimental solubility on 20,000 compounds bins solubility into three ranges; a low range of 10 mg/ml or less; a middle range of 15 to 60 $\mu\text{g/ml}$ and a high range of 65 $\mu\text{g/ml}$ or greater. About 80% of the experimental data used in the model building is evenly distributed between the low and high ranges. The model was tested against 10,000 experimental solubility measurements that were not part of the model building data set. About 80% of the test set data was predicted to lie in the low and high solubility bins and the accuracy of the prediction was 80%. The calculated solubility for the 47,680 combinatorial compounds differed markedly by synthesis protocol with significant populations of protocols at both extremes of solubility. This finding differs markedly from that of permeability. Solubility in combinatorial libraries depends very much on the synthesis protocol and it is very possible for some synthesis protocols to result in poorly aqueous soluble compounds. This and other examples I have investigated leads me to the conclusion that in general poor permeability is not a problem in combinatorial libraries but poor aqueous solubility is indeed a common problem.

The title of this article includes mention of “people issues”. A specific example of a “people issue” is found in the area of experimental solubility profiling of combinatorial libraries. The question relevant to people issues is this. Is it possible to improve the solubility profile of a combinatorial library by incorporating experimental solubility feedback from early exemplars of a library? Understanding this question and how it relates to people issues requires an understanding of the stages involved in the experimental production of a combinatorial library. The experimental component of combinatorial library production typically involves two stages.

These are a protocol development stage followed by a production stage. In the protocol development stage the chemistry to translate the computational design into chemical reality is explored and optimized. Reaction conditions are explored and optimized. Steric and electronic boundaries for reaction components giving acceptable yields are defined and the reaction schemes are converted into formats suitable for robotic implementation. Invariably the protocol development step is the experimental rate determining step. Protocol development is much slower than library production. That is, it takes much longer to work out the chemistry than it does to actually make the compounds once the chemistry is worked out. Compounds first become available for experimental solubility testing in the protocol development stage. The critical issue is timing. How early can the compounds be obtained? The related people issue is this. How early must experimental data be obtained in order for people to change their behavior? Our experience has been that people (chemists) are not willing to change behavior if experimental feedback on solubility comes late in the rate determining step of combinatorial library construction. This is a people factor. When people have invested significant time in protocol development they are very unwilling to change their plans based on late developing experimental data. So finding out that there are likely severe solubility flaws in a library design late in the protocol development stage does not have value in terms of changing the library properties. The chemist has performed most of the work in the rate determining step and is unwilling to stop or radically change chemistry late in the process. The library goes into production regardless of the solubility profile of the exemplars if the feedback occurs late in protocol development. This problem is not easily solved. It is very difficult to obtain exemplars which span the chemical space of the protocol design at an early stage. How early would exemplars have to be obtained so that people factors would permit the experimental solubility data to make a difference? My guess is that it would have to occur in the first 10-15% of the protocol development stage to make a difference. The people factors are very strong. Once chemists strongly commit themselves to protocol development they are very reluctant to stop or even to make very radical changes.

The details of the causes of poor aqueous solubility are important in terms of understanding possible solutions. In a simplistic sense poor aqueous solubility can be thought of as arising from some combination of two distinctly different causes. The balance of these causes differs from compound to compound. At one extreme one cause of poor aqueous solubility lies in what can be termed the cavity making problem. For a compound to dissolve, a cavity (a hole) has to

be made in water. Strong hydrogen bonds must be disrupted to make the cavity. This costs a lot in terms of energy. Some of this energy can be regained if the drug forms favorable interactions with water once it is placed in the hole. However if the compound is very lipophilic few favorable interaction will be formed with water. Hence a large lipophilic compound will be very insoluble in water. A large lipophilic compound requires a large cavity and forming the large cavity costs a lot energetically in terms of many broken hydrogen bonds and little of the energy cost will be reclaimed because there will be few favorable interactions between the lipophilic compound and water. This extreme of solubility is relatively easy to predict computationally. For example 75% of compounds whose lipophilicity exceeds the “rule of 5” limit of $\log P = 5$ have poor aqueous solubility of less than 20 $\mu\text{g/ml}$ in our discovery turbidimetric solubility assay.

At the other extreme one cause of poor aqueous solubility lies in what can be termed the crystal packing problem. A crystalline compound must be liberated from its crystal lattice before it can dissolve. The strongest crystal lattice interactions arise from intermolecular hydrogen bond interactions and packing interactions between the compound and its neighbors in the adjoining unit cells. These interactions can be visualized in the crystal packing diagram that can be generated from a single crystal x-ray. Melting point is the single simple property that is most useful in terms of characterizing crystal packing interactions. A high melting point is indicative of strong intermolecular crystal packing interactions and a high melting point compound is likely to have poor aqueous solubility. A rule of thumb is that a hundred degree increase in melting point decreases aqueous solubility by a factor of ten. This rule of thumb only holds for neutral compounds. The solubility of organic compound salts is not predicted by melting point. Unfortunately for the prediction of aqueous solubility, melting point data is completely absent from combinatorial libraries. Combinatorial compounds purified by automated methods are not subjected to crystallization and are usually isolated in amorphous form. No melting point data exists for these compounds. Clearly it would be very advantageous to have some type of computational prediction method for melting point as an indication of compounds likely to be insoluble because of crystal packing interactions. Unfortunately, reliable methods to predict melting point do not currently exist. Among compounds intended as drugs poor solubility due to strong crystal packing is common. For example in our extensive turbidimetric solubility testing over one half of compounds found to be experimentally insoluble were not excessively large or lipophilic. The prediction of poor solubility due to crystal packing remains a major

unmet computational need in drug discovery because of the relevance to poor aqueous solubility.

CONCLUSION

Oral absorption depends on adequate solubility and intestinal permeability. A compound is insoluble because it is either too lipophilic or the inter molecular crystal packing forces for the compound are too strong. Globally, in the current era, poor aqueous solubility is the single largest physicochemical problem hindering drug oral activity. Among combinatorial libraries, poor solubility is a frequently encountered problem but poor permeability is seldom a problem. The relative importance of poor solubility vs. poor permeability as a source of poor oral activity is very dependent on the method by which leads are generated as can be seen by an examination of the time dependent trends in Merck vs. Pfizer, Groton clinical candidates. Dealing with solubility or permeability problems in an early discovery setting is not purely a technical issue of assay design or computational prediction. People and organizational issues are extremely important. Assay or computational results must be communicated to medicinal chemists in a manner that allows chemists to decide how to modify chemical structure. Communication with chemists is best when it takes advantage of the chemists' superb pattern recognition skills and is least effective when presented in an equation format or in terms that cannot be equated with chemical structure.

REFERENCES AND NOTES

- [1] Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Del. Rev.* **23**:3-25.
- [2] Lipinski, C. A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharm. Tox. Meth.* **44**:235-249.