The Influence of Target Family and Functional Activity on the Physicochemical Properties of Pre-Clinical Compounds

Richard Morphy[†]

Medicinal Chemistry Department, Organon Laboratories, Newhouse, Lanarkshire ML1 5SH, U. K.

Received December 5, 2005

The target families of greatest interest in drug discovery can be differentiated on the basis of the physicochemical properties of their pre-clinical ligands. The ligands for peptidergic targets, such as peptide GPCRs and integrin receptors, possess significantly higher median property values than those for aminergic targets, such as monoamine transporters and GPCRs. The ligands for peptide GPCRs were found to be less efficient, in terms of their binding energy per unit of molecular weight or lipophilicity, than ligands for monoamine GPCRs. The changes in the property values during the optimization process were found to vary only slightly across the target families, with the main determinant of the drug-likeness of the optimized compounds being the profile of the starting compounds. Agonists for monoamine GPCRs, opioid receptors and ion channels were typically smaller and less lipophilic than the antagonists, but there was no difference between the agonists and the antagonists for peptide GPCRs and nuclear receptors.

Introduction

The influence of physicochemical properties on the pharmacokinetic behavior of drug molecules has been the subject of intense interest over the past few years since the publication of Lipinski's seminal work on the rule-of-5 (RO5) in 1997.[1] This was followed by the work of Teague et al.[2] and Hann et al.,[3] which highlighted the fact that molecules tend to increase in MW and cLogP during optimization. More recent work has examined the influence of the degree of (pre-)clinical advancement of molecules on physicochemical properties as well as the influence of the disease area,[4] launch date,[5] and route of administration.[6]

Over recent years, there has been an increasing amount of anecdotal evidence indicating that the discovery of orally active ligands for some targets and target families is more challenging than that of others. The aim of this work was to study the influence of target family on the physicochemical properties of pre-clinical ligands and thereby provide, for the first time, some quantitative data supporting these perceived differences. The influence of functional activity was also explored.

Over the past few years, a database (called SCOPE) has been assembled at Organon, which consists of a large number of optimizations extracted predominantly from the primary literature. For each selected publication, the structures of the both the starting compound and the most highly optimized compound have been extracted.

Data Source. Currently, the SCOPE database contains a total of 1860 optimizations, 1630 (88%) from the literature, and 230 (12%) from internal Organon projects. Each entry was annotated by target family, and this feature allowed a detailed analysis of the influence of these features on the physicochemical properties of the optimized compounds and the changes in those properties during optimization. The distribution of the major target families within the database is shown in Figure 1. For reasons of statistical validity, only target families that represented 2% or more of the total database were considered in this analysis, representing in total 89% of the database. These families included those of greatest current interest in drug discovery. Many entries also contained information about binding affinity and functional activity, and this allowed a further analysis of the relationship between physicochemical and biological properties.

The database contains predominantly pre-clinical compounds, although a very small number of compounds that reached the market are included (8 in total). The majority of the entries are from the year 2000 onward when a systematic program of abstraction of four major medicinal chemistry journals began (*Bioorganic & Medicinal Chemistry*, *Bioorganic & Medicinal Chemistry Letters*, *European Journal of Medicinal Chemistry*, and *Journal of Medicinal Chemistry*). A smaller number of entries from the 1990−1999 period were also added. The year by year distribution of entries is shown in Figure 2. SCOPE is not a comprehensive database of all optimized compounds from the most recent literature. Because its principal aim is to capture information about the optimization process itself, only publications containing clearly identified starting and optimized compounds are abstracted. The abstraction policy for the SCOPE database is to select, as the optimized compound, the compound that was subjected to the most rigorous testing, and this is usually highlighted in the publication abstract or conclusion. It is not necessarily the most potent compound in the primary in vitro assay but rather the compound with the most rounded properties overall in terms of, for example, in vitro and in vivo potency, selectivity, and pharmacokinetic properties.

Property Calculations and Statistical Analysis. For each optimized compound from both the full SCOPE set, the target family subsets, and the functional activity subsets, six physicochemical properties were calculated: molecular weight (MW), cLogP, polar surface area (PSA),[7] the number of hydrogen bond acceptors (HBA), the number of hydrogen bond donors (HBD), and the number of rotatable bonds (RB).[8] Because the database also contained the starting compound in each case, the changes in these same properties during optimization could also be studied.

At the start of this work, it became apparent that for some of the target family and functional subsets, the property distributions for the optimized compounds and trajectories were not

[†] To whom correspondence should be addressed. Phone: +44 (0)1698 736000. Fax: +44 (0)1698 736187. E-mail: r.morphy@organon.co.uk.

10.1021/jm0512185 CCC: $33.50
Published on Web 04/13/2006

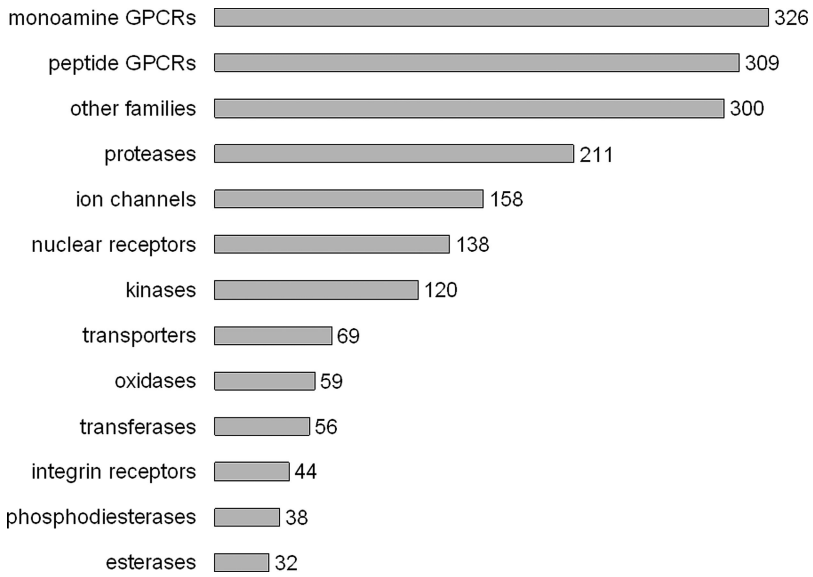

Figure 1. Target family distribution of SCOPE database entries.

Table 1. Molecular Weight (MW) Data Classified by Target Family

target family subset	number of entries	starting compd median (mean)	optimized compd median (mean)	optimized compd lower quartile	optimized compd upper quartile	change in property median (mean)	95% CI (change in property)[a]	p value (change in property)[b]
full SCOPE set	1680	382 (393)	422 (435)	353	504	30 (42)	(32, 39)	0
esterases	32	349 (336)	383 (412)	339	435	51 (76)	(35, 95)	0
GPCRs (all)[c]	755	391 (402)	433 (440)	358	517	30 (38)	(30, 39)	0
GPCRs (monoamine)	326	347 (347)	375 (377)	295	444	24 (30)	(21, 33)	0
GPCRs (peptide)	309	451 (465)	510 (513)	434	581	44 (48)	(39, 54)	0
integrin receptor	41	454 (449)	466 (496)	436	575	33 (47)	(12, 78)	0.006
ion channels	158	311 (328)	364 (373)	300	429	33 (45)	(29, 47)	0
kinases	120	349 (360)	392 (406)	343	470	33 (46)	(29, 56)	0
nuclear receptors	138	410 (406)	421 (431)	381	492	18 (25)	(12, 33)	0
oxidases	59	314 (314)	357 (355)	315	403	21 (41)	(16, 56)	0
phosphodiesterases	38	415 (409)	462 (465)	392	569	39 (56)	(28, 79)	0
proteases	211	421 (427)	467 (468)	409	524	32 (41)	(29, 49)	0
transferases	56	451 (502)	521 (539)	392	693	30 (38)	(19, 50)	0
transporters	69	299 (306)	325 (335)	270	395	25 (29)	(15, 36)	0

[a] The 1-sample Wilcoxon 95% confidence interval. [b] The *p* value from 1-sample Wilcoxon signed rank test of the median = 0 versus median not = 0. [c] The GPCR (monoamine) and GPCR (peptide) subsets are contained within the GPCR (all) subset.

normally distributed but were skewed and often included extreme outliers. For reasons of consistency, all of the data sets were analyzed using nonparametric rank statistical methods rather than parametric t-tests. The Wilcoxon signed rank test was used to examine the significance of the changes in the properties during optimization, and the Mann−Whitney rank test was used to explore the significance of the differences in properties between the target family subsets. For that reason, more emphasis on the analysis and interpretation was placed upon the median values, although the mean values are also quoted for reference purposes.

Results

Molecular Weight. The median MW for the full set of 1680 optimized compounds was 422, and the mean was 435 (Table 1). These values are notably higher than the reported values for marketed oral drugs. For example, Vieth et al. reported a median MW of 322 and a mean MW of 344 for 1202 oral drugs.[6] Vieth et al.[6] and Blake et al.[9] reported median MWs of 415 (mean of 448) and 393, respectively, for a range of preclinical compounds, and these figures are consistent with the SCOPE averages. Among the target family subsets, the median MW was highest for the peptide GPCR ligands and the transferase inhibitors with values of 510 and 521, respectively (Table 1). At the other extreme, the ligands for transporters had a median MW of just 325.

The median increase in MW during optimization was 30 (mean 42). This was lower than the median MW increase of 70 reported by Oprea et al.[10] during the optimization of leads to drugs. However, the SCOPE mean increase is identical to the reported mean MW increase of 42 reported by Hann et al.[3] for the lead to drug process. There is a statistically significant increase in MW for all of the target family subsets (p value < 0.001), indicating that this underlying trend is consistent and highly pronounced during the process of optimization. The biggest increases in MW were found for the peptide GPCR and esterase groups, 44 and 51, respectively. The lowest MW change was found for the nuclear receptor ligands (21).

The target family subsets were divided into clusters by determining whether the MWs of the optimized compounds showed statistically significant differences from each other (p value < 0.1).[11] The MWs of the peptide GPCR, transferase, and integrin receptor optimized compounds were not significantly different but were all significantly higher than the MWs for the protease and phosphodiesterase (PDE) optimized compounds. In turn, the MWs for the protease and PDE subsets

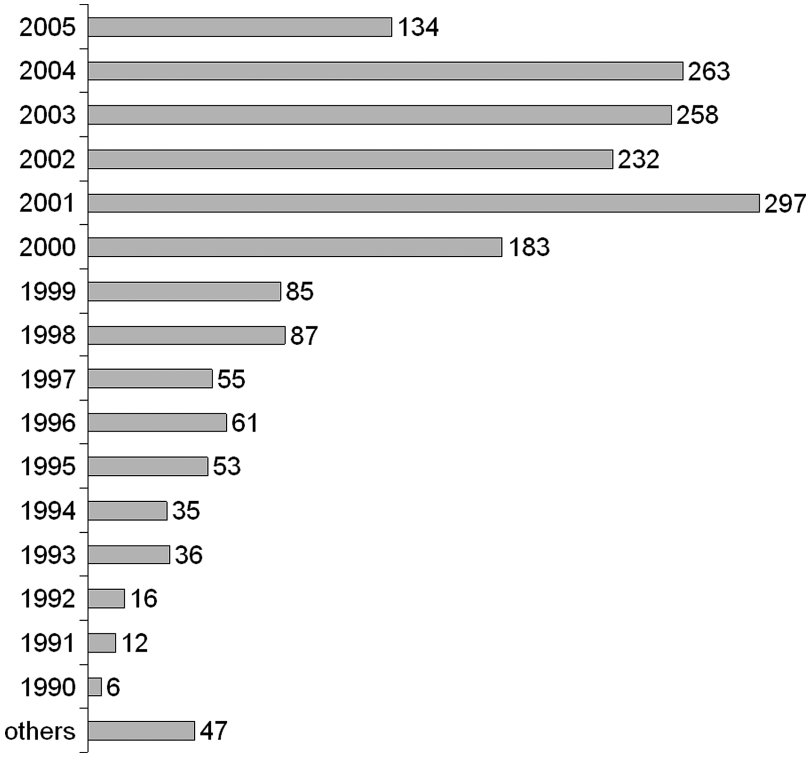

Figure 2. Year-by-year distribution of SCOPE database entries.

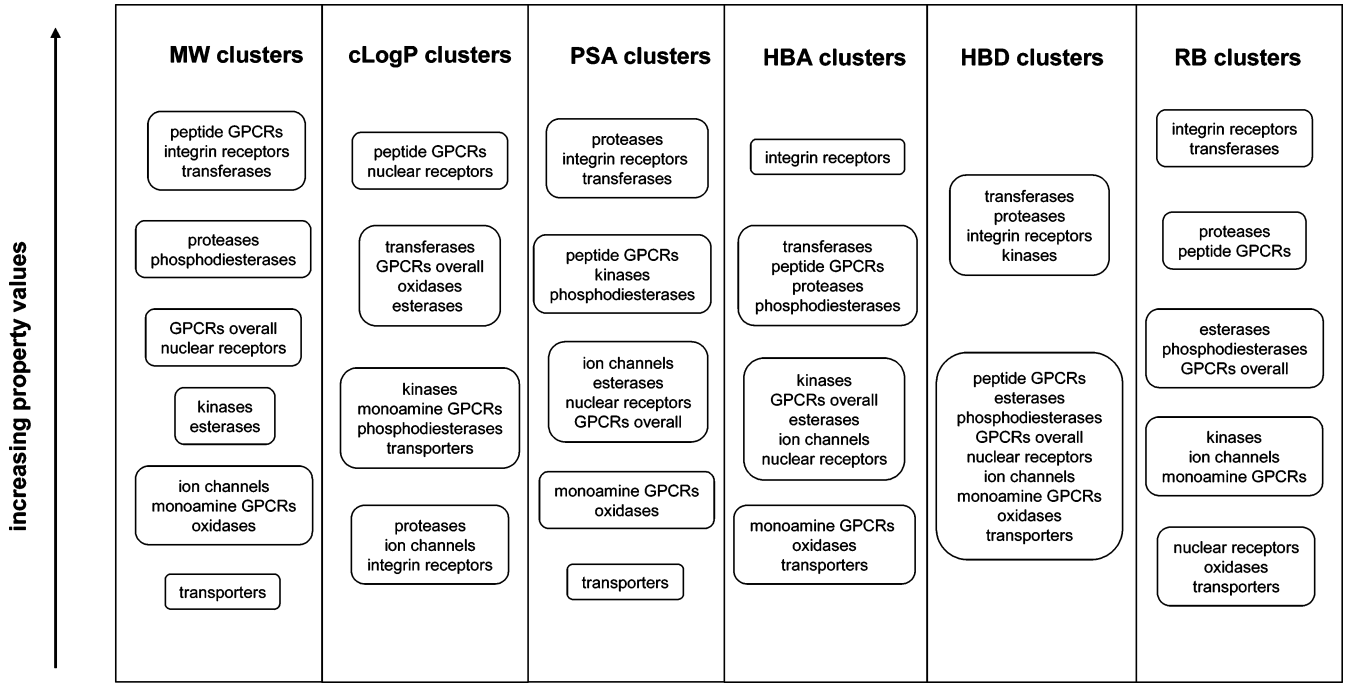

Figure 3. Classification of target families on the basis of six physicochemical properties for their optimized ligands (MW, cLogP, polar surface area (PSA), the number of hydrogen bond acceptors (HBA), the number of hydrogen bond donors (HBD), and the number of rotatable bonds (RB)).

were not significantly different from each other but were significantly higher than those for the nuclear receptors. Similarly, the MWs for the kinase and esterase subsets were not different from each other but were significantly higher than those for the monoamine GPCRs, ion channels, and oxidases. The transporters had the lowest median MW ligands of all. Overall, this process of cross-comparisons using the Mann−Whitney test created a sequence of six statistically defined MW clusters (Figure 3).

The higher molecular weights of the optimized compounds from the uppermost cluster could be due to a higher MW of the starting compounds, a higher increase during optimization, or a combination of the two. When the MWs for peptide GPCRs, integrin receptors, proteases, and transferases were compared with those for monoamine GPCRs, transporters, and oxidases, it was clear that for the former the starting compounds were much larger, and, in addition, for peptide GPCRs and integrin receptors, there was a greater increase in size during optimization (Table 1).

For the full SCOPE database, the receptor antagonists had the highest median MW, followed by the enzyme inhibitors, the transporter inhibitors, and then the receptor agonists (Table

Table 2. Physicochemical Property Data for the Optimized Compounds Classified by Target Family

target family subset	function	number of entries[a]	MW median (mean)	cLogP median (mean)	PSA median (mean)	HBA median (mean)	HBD median (mean)	RB median (mean)
full set	agonist	314	**402** (405)	4.1 (4.2)	48 (53)	**3** (4)	1 (2)	**5** (6)
	antagonist	682	**437** (448)	4.4 (4.3)	55 (60)	**4** (4)	1 (2)	**6** (7)
	inhibitor	675	412 (435)	3.4 (3.4)	66 (74)	5 (5)	2 (2)	6 (7)
GPCRs (all)	agonist	118	**380** (377)	**3.8** (3.8)	**39** (46)	**3** (3)	1 (2)	6 (6)
	antagonist	380	**445** (458)	**4.5** (4.5)	**53** (57)	**4** (4)	1 (1)	6 (7)
GPCRs (monoamine)	agonist	61	**269** (329)	**3.3** (3.2)	36 (41)	3 (3)	**2** (2)	4 (5)
	antagonist	148	**394** (395)	**4.2** (4.0)	41 (43)	3 (3)	**1** (1)	5 (5)
GPCRs (peptide)	agonist[b]	38	**446** (455)	4.9 (4.8)	**42** (48)	**4** (4)	1 (1)	**6** (7)
	agonist[c]	12	548 (520)	5.6 (5.8)	64 (61)	4 (4)	1 (2)	9 (8)
	antagonist	165	**523** (535)	5.0 (5.1)	**65** (69)	**5** (5)	1 (2)	**8** (9)
ion channels	agonist	25	**285** (312)	**2.1** (2.2)	**45** (47)	3 (4)	**1** (1)	**3** (5)
	antagonist	94	**369** (388)	**3.2** (3.1)	**57** (64)	4 (4)	**2** (2)	**5** (6)
nuclear receptors	agonist	109	419 (425)	4.9 (5.3)	50 (53)	3 (4)	1 (1)	4 (5)
	antagonist	22	406 (429)	4.7 (5.0)	53 (53)	5 (4)	1 (1)	5 (6)

[a] The dataset used for the functional activity analysis was different from that used for the target family analysis because a functional activity assignment was not available for all SCOPE entries. [b] Full peptide GPCR agonist subset containing opioid ligands. [c] Peptide GPCR agonist subset minus opioid ligands. Bold text indicates Mann−Whitney $p < 0.05$ for the comparison between the agonist and antagonist subsets for each target family group.

Table 3. cLogP Data Classified by Target Family

target family subset	number of entries	starting compd median (mean)	optimized compd median (mean)	optimized compd lower quartile	optimized compd upper quartile	change in property median (mean)	95% CI (change in property)[a]	p value (change in property)[b]
full SCOPE set	1680	3.7 (3.6)	4 (4)	2.8	5.2	0.2 (0.4)	(0.2, 0.4)	**0**
esterases	32	3.3 (3.5)	4.2 (4.3)	2.8	5.4	0.7 (0.8)	(−0.1, 1.5)	0.094
GPCRs (all)[c]	755	4.1 (4)	4.4 (4.3)	3.3	5.5	0.3 (0.3)	(0.2, 0.4)	**0**
GPCRs (monoamine)	326	3.5 (3.3)	3.8 (3.7)	2.9	4.7	0.2 (0.4)	(0.2, 0.5)	**0**
GPCRs (peptide)	309	4.8 (4.8)	5 (5)	3.8	5.9	0.3 (0.2)	(0, 0.4)	**0.042**
integrin receptor	41	1.8 (2.2)	3.1 (2.9)	0.7	4.5	0.4 (0.7)	(0, 1)	0.057
ion channels	158	2.9 (2.7)	3 (3)	1.6	4.7	0.2 (0.3)	(−0.1, 0.5)	0.112
kinases	120	3.4 (3.1)	3.8 (3.5)	2.1	4.7	0.1 (0.4)	(0, 0.5)	0.078
nuclear receptors	138	4.9 (5.2)	5.1 (5.3)	3.9	6.5	−0.1 (0.1)	(−0.2, 0.3)	0.679
oxidases	59	3.7 (3.6)	4.2 (4.2)	2.9	5.4	0.4 (0.6)	(0.1, 1)	**0.008**
phosphodiesterases	38	3.1 (3.1)	3.5 (3.6)	2.9	4.2	0.3 (0.5)	(0, 0.9)	0.061
proteases	211	2.8 (2.7)	3.2 (3)	1.7	4.1	0.3 (0.3)	(0.1, 0.5)	**0.006**
transferases	56	4 (3.9)	4.5 (4.3)	2.9	5.3	0.4 (0.4)	(0, 0.7)	0.069
transporters	69	3.2 (3.4)	3.7 (4)	3.0	5.1	0.4 (0.6)	(0.2, 0.8)	**0.001**

[a] The 1-sample Wilcoxon 95% confidence interval. [b] The p value from 1-sample Wilcoxon signed rank test of the median = 0 versus median not = 0. [c] The GPCR (monoamine) and GPCR (peptide) subsets are contained within the GPCR (all) subset.

2). For the monoamine GPCR optimized ligands, there was a particularly pronounced difference in MW between the agonists and antagonists. There was also a difference between the agonists and antagonists for the peptide GPCR optimized ligands, but this was found to be due to a large number of opioid receptor agonists in the agonist subset. When the latter were excluded, there was no significant difference between the MW of the peptide GPCR agonists and antagonists.

cLogP. The median cLogP for the full set of optimized compounds was 3.7, and the mean was 3.6 (Table 3). These figures were significantly higher than the medians of 2.3 and 2.5 reported by Vieth et al.[6] and Blake et al.,[9] respectively for oral drugs but were very similar to those reported for pre-clinical compounds by the same researchers (medians of 3.5 and 3.23, respectively). The target family subsets with the highest median cLogP values for the optimized compounds were the nuclear receptor ligands (5.1) and the peptide GPCR ligands (5.0), whereas the integrin receptor ligands were much less lipophilic (3.1).

The median increase in cLogP during optimization for the full set was 0.2, and the mean was 0.4. For comparison, Hann et al.[3] reported a mean increase of 0.5, and Oprea et al.[10] reported a median increase in cLogP of 0.67, during the optimization of leads to drugs. The extent of the increase in cLogP during optimization was highly conserved across all the target families, typically being in the range of 0.2 to 0.4. The exception was the nuclear receptor family for which there was no increase during optimization, but in this case the starting compounds were already highly lipophilic.

Using the Mann−Whitney test, a series of cross-comparisons between the target families allowed a classification on the basis of the median cLogP differences and the p values (Figure 3).[11] The most lipophilic cluster contained the peptide GPCRs and the nuclear receptors, and the least lipophilic cluster contained the proteases, ion channels, and integrin receptors.

There was no significant difference in lipophilicity between the receptor agonists and the antagonists present in the full SCOPE database (Table 2). However, the receptor agonists were significantly less lipophilic than the antagonists for the full GPCR, the monoamine GPCR, and ion channel subsets. The enzyme and transporter inhibitors were close to an order of magnitude less lipophilic than the receptor agonists or antagonists.

PSA. The median PSA for the full set of optimized compounds was 58 Å^2, and the mean was 63 Å^2 (Table 4). In contrast, the mean PSA values reported by Vieth et al.[6] were 78 Å^2 for oral drugs and 96.7 Å^2 for pre-clinical compounds. Median values were not reported. The median PSA value reported by Blake et al.[9] was 122 Å^2 for oral drugs and 137 Å^2

Table 4. Polar Surface Area (PSA) Data Classified by Target Family

target family subset	number of entries	starting compd median (mean)	optimized compd median (mean)	optimized compd lower quartile	optimized compd upper quartile	change in property median (mean)	95% CI (change in property)[a]	p value (change in property)[b]
full SCOPE set	1680	52 (58)	58 (63)	38	80	1 (5)	(4, 5)	**0**
esterases	32	41 (50)	48 (59)	36	75	6 (8)	(1, 14)	**0.012**
GPCRs (all)[c]	755	45 (52)	51 (56)	35	72	0 (4)	(3, 5)	**0**
GPCRs (monoamine)	326	37 (42)	41 (43)	29	55	0 (1)	(0, 4)	**0.011**
GPCRs (peptide)	309	53 (58)	62 (65)	43	83	6 (7)	(5, 8)	**0**
integrin receptor	41	90 (95)	89 (95)	76	116	6 (0)	(−8, 16)	0.499
ion channels	158	47 (53)	54 (59)	39	73	4 (6)	(3, 8)	**0**
kinases	120	59 (66)	65 (71)	55	78	5 (5)	(2, 9)	**0.001**
nuclear receptors	138	50 (48)	52 (54)	33	67	3 (6)	(2, 8)	**0**
oxidases	59	45 (42)	43 (46)	30	60	0 (4)	(0, 8)	**0.079**
phosphodiesterases	38	64 (62)	64 (69)	56	80	1 (7)	(0, 13)	**0.015**
proteases	211	82 (85)	89 (90)	73	107	2 (5)	(2, 7)	**0.001**
transferases	56	71 (84)	83 (90)	56	123	1 (6)	(0, 9)	**0.024**
transporters	69	19 (27)	22 (28)	17	40	0 (1)	(0, 4)	0.428

[a] The 1-sample Wilcoxon 95% confidence interval. [b] The *p* value from 1-sample Wilcoxon signed rank test of the median = 0 versus median not = 0. [c] The GPCR (monoamine) and GPCR (peptide) subsets are contained within the GPCR (all) subset.

Table 5. Hydrogen Bond Acceptor (HBA) Data Classified by Target Family

target family subset	number of entries	starting compd median (mean)	optimized compd median (mean)	optimized compd lower quartile	optimized compd upper quartile	change in property median (mean)	95% CI (change in property)[a]	p value (change in property)[b]
full SCOPE set	1680	4 (4)	4 (4.4)	3	6	0 (0.5)	(0.5, 0.5)	**0**
esterases	32	3 (3.7)	4 (4.3)	3	6	0 (0.6)	(0, 1)	**0.014**
GPCRs (all)[c]	755	3 (3.8)	4 (4.2)	3	5	0 (0.4)	(0.5, 0.5)	**0**
GPCRs (monoamine)	326	3 (3.2)	3 (3.4)	2	4	0 (0.2)	(0, 0.5)	**0.003**
GPCRs (peptide)	309	4 (4.2)	5 (4.8)	3	6	0 (0.6)	(0.5, 0.5)	**0**
integrin receptor	41	6 (5.8)	6 (6.2)	5	8	0.5 (0.5)	(0, 1)	0.173
ion channels	158	3 (3.7)	4 (4.2)	3	5	0 (0.5)	(0, 0.5)	**0**
kinases	120	4 (4.1)	4 (4.6)	3	5	0 (0.5)	(0.5, 1)	**0**
nuclear receptors	138	3 (3.4)	3 (3.8)	2	5	0 (0.4)	(0, 0.5)	**0.001**
oxidases	59	3 (2.8)	3 (3.3)	2	4	0 (0.5)	(0, 0.5)	**0.01**
phosphodiesterases	38	5 (4.9)	5 (5.2)	4	6	0 (0.3)	(0, 1)	0.27
proteases	211	5 (4.9)	5 (5.4)	4	6	0 (0.6)	(0.5, 0.5)	**0**
transferases	56	4 (5.3)	5 (5.7)	4	8	0 (0.4)	(0, 1)	**0.029**
transporters	69	3 (2.7)	3 (2.6)	2	3	0 (−0.1)	(0, 0)	0.361

[a] The 1-sample Wilcoxon 95% confidence interval. [b] The *p* value from 1-sample Wilcoxon signed rank test of the median = 0 versus median not = 0. [c] The GPCR (monoamine) and GPCR (peptide) subsets are contained within the GPCR (all) subset.

for pre-clinical compounds. The large discrepancy between different research groups almost certainly reflects differences in the methods of the calculation of PSA rather more than any inherent differences in the properties of the compounds surveyed. The highest medians were found for the integrin receptor ligands and the protease inhibitors (89 Å^2). The lowest median (22 Å^2) was observed for the transporter ligands.

For the full set of optimized compounds, there was very little change in PSA during optimization, with a median increase of just 1 Å^2. Across the target family subsets, the increases in PSA during optimization were consistently small. The highest statistically significant increases were found for the peptide GPCR and esterase ligands (6 Å^2).

The target families were classified on the basis of the PSA values and the Mann−Whitney *p* values using the procedure described earlier (Figure 3). The transferase, protease, and integrin receptor ligands had the highest values and the transporter ligands by far the lowest.

The median PSA for the agonists present in the full SCOPE set of optimized compounds was lower than that for the antagonists or inhibitors (Table 2).

HBA. The median number of HBAs for the full set was 4, and the mean was 4.4 (Table 5). This was identical to the median value reported by Blake et al.[9] for pre-clinical compounds. Vieth et al.[6] reported a median of 3 for both oral drugs and pre-clinical compounds. The highest median was found for the integrin receptor ligands (6) and the lowest for the transporter ligands (3). Only the integrin receptor ligands showed a statistically significant increase in the number of HBAs (0.5). Integrin receptor ligands had the greatest number of HBAs and monoamine GPCR, oxidase, and transporter ligands the fewest (Figure 3).

The inhibitors had the highest median number of HBAs followed by the antagonists and then the agonists (Table 2). For the full GPCR and peptide GPCR subsets, the median number of HBAs in the agonists was lower than that of the antagonists.

HBDs. The full set of optimized compounds showed a median number of 1 HBD and a mean of 1.9 (Table 6). This was virtually identical to the figures reported by Vieth et al.[6] for oral drugs (median 1; mean 1.8) but somewhat less than Vieth's and Blake's[9] figures for pre-clinical compounds (median 2; mean 2.1). The integrin receptor ligands and the protease inhibitors possessed the highest number of HBDs with a median of 3 (Figure 3). There was no increase in the number of HBDs during optimization for any of the target family subsets.

For the peptide GPCR and ion channel subsets, the median number of HBDs in the agonists was lower than that of the antagonists, whereas for the monoamine GPCRs, the reverse was true (Table 2). When the opioid ligands were excluded,

Table 6. Hydrogen Bond Donor (HBD) Data Classified by Target Family

target family subset	number of entries	starting compd median (mean)	optimized compd median (mean)	optimized compd lower quartile	optimized compd upper quartile	change in property median (mean)	95% CI (change in property)[a]	p value (change in property)[b]
full SCOPE set	1680	1 (1.9)	1 (1.9)	1	3	0 (0)	(0, 0)	0.892
esterases	32	1 (1.5)	1.5 (1.7)	1	2	0 (0.2)	(0, 0.5)	0.363
GPCRs (all)[c]	755	1 (1.5)	1 (1.5)	1	2	0 (0)	(0, 0)	0.496
GPCRs (monoamine)	326	1 (1.3)	1 (1.2)	0	2	0 (−0.1)	(0, 0)	0.143
GPCRs (peptide)	309	1 (1.6)	1 (1.7)	1	2	0 (0.1)	(0, 0)	0.277
integrin receptor	41	3 (3.6)	2 (3.0)	2	4.75	0 (−0.6)	(−1, 0)	0.139
ion channels	158	2 (1.7)	2 (1.9)	1	3	0 (0.2)	(0, 0.5)	**0.016**
kinases	120	2 (2.6)	2 (2.6)	2	3	0 (0)	(0, 0)	1
nuclear receptors	138	1 (1.2)	1 (1.2)	1	2	0 (0)	(0, 0)	0.873
oxidases	59	1 (1.1)	1 (1)	0	2	0 (−0.1)	(0, 0)	0.411
phosphodiesterases	38	1 (1)	1 (1.3)	1	2	0 (0.2)	(0, 0.5)	0.112
proteases	211	3 (3.4)	3 (3.3)	2	5	0 (−0.1)	(0, 0)	0.216
transferases	56	2.5 (2.9)	2 (2.9)	1	4.75	0 (0)	(−0.5, 0)	0.909
transporters	69	1 (0.7)	1 (0.9)	0	1	0 (0.2)	(0, 0)	**0.018**

[a] The 1-sample Wilcoxon 95% confidence interval. [b] The *p* value from 1-sample Wilcoxon signed rank test of the median = 0 versus median not = 0. [c] The GPCR (monoamine) and GPCR (peptide) subsets are contained within the GPCR (all) subset.

Table 7. Rotatable Bond (RB) Data Classified by Target Family

target family subset	number of entries	starting compd median (mean)	optimized compd median (mean)	optimized compd lower quartile	optimized compd upper quartile	change in property median (mean)	95% CI (change in property)[a]	p value (change in property)[b]
full SCOPE set	1680	5 (6)	6 (6.7)	4	9	0 (0.7)	(0.5, 0.5)	**0**
esterases	32	5.5 (5.9)	6 (6.5)	4	8.75	0.5 (0.6)	(−0.5, 2)	0.213
GPCRs (all)[c]	755	5 (5.9)	6 (6.6)	4	9	0 (0.6)	(0.5, 0.5)	**0**
GPCRs (monoamine)	326	5 (4.9)	5 (5.2)	3	7	0 (0.3)	(0, 0.5)	**0.018**
GPCRs (peptide)	309	6 (7)	7 (7.9)	6	10	1 (0.9)	(0.5, 1)	**0**
integrin receptor	41	9 (10.1)	9 (10.2)	7	11	0.5 (0)	(−1, 1.5)	0.850
ion channels	158	4 (4.7)	4 (5.7)	3	7	0 (1)	(0.5, 1)	**0**
kinases	120	4 (4.9)	5 (5.5)	4	6	1 (0.6)	(0.5, 1)	**0.004**
nuclear receptors	138	4 (4.5)	4 (5.3)	2	8	0 (0.8)	(0, 1)	**0**
oxidases	59	3 (3.1)	4 (4.2)	3	5	0 (1.1)	(0.5, 1.5)	**0**
phosphodiesterases	38	7 (6.3)	7 (6.7)	4.75	9	0 (0.4)	(0, 1)	0.136
proteases	211	7 (7.8)	8 (8.2)	6	10	0 (0.4)	(0, 0.5)	**0.007**
transferases	56	10 (10.8)	10 (11.7)	7	19	1 (0.9)	(0.5, 1.5)	**0.005**
transporters	69	4 (4.6)	3 (4.5)	2	7.5	0 (−0.1)	(−0.5, 0)	0.291

[a] The 1-sample Wilcoxon 95% confidence interval. [b] The *p* value from 1-sample Wilcoxon signed rank test of the median = 0 versus median not = 0. [c] The GPCR (monoamine) and GPCR (peptide) subsets are contained within the GPCR (all) subset.

there was no difference between peptide GPCR agonists and antagonists.

Rotatable Bonds (RB). For the full set of optimized compounds, there was a median number of six rotatable bonds and a mean of 6.7 (Table 7). This was very similar to the median figures reported reported by Vieth et al.[6] and Blake et al.[9] for oral drugs (5 and 6, respectively) and broadly comparable to the median of 7 reported by the same researchers for preclinical compounds. The most flexible ligands were those for transferases and integrin receptors, whereas transporter, nuclear receptor, and oxidase ligands were the most rigid (Figure 3).

For the full set, the median number of rotatable bonds was lower for the agonists than that for the antagonists or inhibitors (Table 2). For the full GPCR, peptide GPCR, and ion channel subsets, the agonists were more rigid than the antagonists. However, when the opioid ligands were excluded from the peptide GPCR subset, there was also no difference between the agonists and antagonists. For the monoamine GPCRs and nuclear receptors, there was also no significant difference.

Discussion

Target Family Relationships. It is clear from this study that the major SCOPE target families that include those of greatest current interest in drug discovery differ considerably in terms of the physicochemical properties of their optimized compounds. There was a high degree of consistency in the rank order of the 13 target family subsets irrespective of what physicochemical property was considered, and it is, therefore, a reasonable proposition to identify an overall rank order on the basis of a consideration of all six physicochemical properties collectively. However, unlike the statistically defined clusters in Figure 3, it is necessary to use an element of judgment to assess the cumulative effect of all six properties.

Peptide GPCRs, integrin receptors, proteases, and transferases gave consistently high property values for the optimized compounds. This group of targets may be regarded as the most challenging in terms of obtaining ligands that combine acceptable biological potency with drug-like physicochemical properties (Figure 4). Several target families consistently gave values for the six physicochemical properties that were in the middle of the range: PDEs, GPCRs (all), kinases, nuclear receptors, and esterases. For these target families, obtaining a drug-like, potent ligand will typically be more feasible than that for the upper group. Four target families exhibited median property values that were consistently low: monoamine GPCRs, ion channels, oxidases, and transporters.

The robustness of the grouping of the target families in Figure 4 can be tested by considering a smaller number of properties.

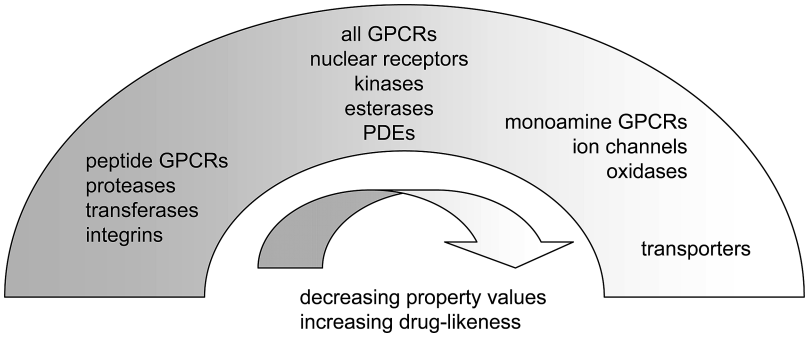

Figure 4. Spectrum of drug-likeness for the 13 target families on the basis of 6 physicochemical properties for their optimized ligands (MW, cLogP, polar surface area (PSA), the number of hydrogen bond acceptors (HBA), the number of hydrogen bond donors (HBD), and the number of rotatable bonds (RB)).

Particular combinations of properties that have been postulated in previous publications to be particularly important for determining the oral absorption of drugs are (1) MW, cLogP, HBD, and HBA;[1] (2) PSA, and RB;[12] or (3) cLogP, HBD, and PSA.[5] These new combinations also address the high correlation and, therefore, possible redundancy that was found between the MW and RB data ($R^2 = 0.83$) and between the PSA and HBA data ($R^2 = 0.92$).[13] When PSA and RB alone are considered, the protease ligands may be regarded as having a similar drug-likeness to transferase, peptide GPCR, and integrin receptor ligands because of their particularly high PSA values. The important point is that the overall rank order of the target families does not seem to substantially change when these smaller sets of properties are considered.

Functional Activity Relationships. The agonists in the SCOPE database were typically smaller and less lipophilic compared to the antagonists, and this effect was particularly pronounced for the monoamine GPCR optimized ligands. This was not unexpected because the medicinal chemistry literature contains many examples of monoamine receptor antagonists that are structurally related to agonists but contain added lipophilic functionality.[14] The vast majority of the optimized ligands for peptide GPCRs in the SCOPE database, and indeed in the general medicinal chemistry literature, are antagonists. The relatively small number of peptide GPCR agonists also appeared at first sight to be smaller than their antagonist counterparts, but this was heavily influenced by the predominance of low MW opioid agonists. The non-opioid peptide GPCR agonists were found to have properties that were very similar to those of the antagonists. In fact, there is some evidence that peptide receptor agonists are often larger because of the need for an extra group that acts as an agonist trigger.[15] The ion channels resembled the monoamine GPCRs in the sense that the agonists had lower MW and lipophilicity than the antagonists, whereas the nuclear receptors resembled the peptide GPCRs with little difference between the two groups.

Efficiency of Ligand Binding. The endogenous ligands for the targets in the upper group (peptide GPCRs, integrin receptors, proteases, and transferases) are predominantly complex peptidic or nucleotidic structures, and the increased size and complexity of the optimized compounds for these target families seem to mirror this. In contrast, targets with less complex endogenous ligands, such as the monoamine GPCRs and transporters, possess less complex optimized compounds. This work is consistent with previous observations that there is a difference between the properties of ligands that bind to aminergic and nonaminergic GPCRs.[16]

There are a number of possible explanations as to why the ligands for aminergic targets are typically smaller and less complex than their peptidergic counterparts. One component may be the approach through which the starting compound was identified by the medicinal chemist. For the monoamine targets, many ligands are clearly analogues of the endogenous ligands such as adrenaline that are small and polar. For peptide targets, high throughput screening (HTS) will have been used more extensively as the source of starting compounds, given that an endogenous ligand-based approach is usually intractable. HTS has been implicated as a cause of higher property values, such as MW and cLogP, in pre-clinical compounds.[17] One reason for this is that HTS collections have historically contained compounds that had high MW and lipophilicity, for example, from early combinatorial chemistry libraries. Optimized compounds in the SCOPE database, for which HTS was known to be the source, do indeed have median and mean MWs (412 and 435) that are higher than those derived from marketed drugs (median 382; mean 390) (Table 8). A similar trend was seen for cLogP, with HTS-derived optimized compounds being more lipophilic. Where HTS was the common source, peptide GPCR ligands were much larger than monoamine GPCR ligands, with median MWs of 421 and 345, respectively. This trend was even more pronounced where a marketed drug was the common source. Peptide GPCR ligands were also more lipophilic than those for monoamine receptors when the source of the starting compound was the same. These trends suggest that the differences between the peptide and monoamine GPCR ligands cannot readily be explained by differences in the source of the starting compounds but rather stem from fundamental differences in the way the ligands are recognized by their respective receptors.

Table 8. MW and cLogP Data for the Optimized Compounds Classified According to the Source of the Starting Compound

source subsets	number of entries[a]	median	mean	*p* value[b]
	MW			
HTS (all targets)	149	412	435	0.0029
drug (all targets)	128	382	390	
HTS (monoamine GPCRs)	17	345	336	0.0049
HTS (peptide GPCRs)	37	421	426	
drug (monoamine GPCRs)	35	312	308	0.0001
drug (peptide GPCRs)	17	464	478	
	cLogP			
HTS (all targets)	149	4.6	4.5	0.0000
drug (all targets)	128	3.7	3.4	
HTS (monoamine GPCRs)	17	3.2	3.8	0.0242
HTS (peptide GPCRs)	37	4.8	4.6	
drug (monoamine GPCRs)	35	3.1	3.0	0.1097
drug (peptide GPCRs)	17	4.2	3.7	

[a] The data set used for the source analysis was different from that used for the target family analysis because the source assignment was not available for all SCOPE entries. [b] Mann−Whitney *p* value. HTS = high throughput screening; drug = marketed drug.

Table 9. Ligand Efficiency Data for Monoamine and Peptide GPCR Ligands

target family subset		number of entries[a]	MW median[b]	clogP median	affinity median (p*Ki*)[c]	affinity per dalton	affinity per log unit of lipophilicity
GPCRs (monoamine)	SC	198	343	3.5	7.77	0.0227	2.25
	OC	198	376	3.8	8.40	0.0223	2.19
GPCRs (peptide)	SC	106	453	5.2	7.78	0.0172	1.51
	OC	106	518	5.0	8.73	0.0169	1.76

[a] The dataset used for the affinity analysis was smaller than that used for the target family analysis in Table 1 because affinity data were not available for all SCOPE entries. [b,c] The differences in the MW and affinity medians between the starting and optimized compounds are significant (Mann−Whitney p value < 0.05). SC = starting compounds; OC = optimized compounds.

This was investigated by exploring the relationships between molecular size, lipophilicity, and binding affinity (p*Ki*). The median affinities of the starting and optimized compounds for the monoamine and peptide GPCRs are shown in Table 9. These data suggest that monoamine GPCR ligands are more efficient than peptide GPCR ligands in terms of their binding energy per unit of MW or cLogP. For MW, monoamine GPCR ligands have an efficiency, as measured by p*Ki* divided by MW, that is approximately 25% higher.

The binding site for the (ant)agonists for Class A monoamine and peptide GPCRs is thought to reside wholly or partly in the transmembrane (TM) regions of the receptors close to the extracellular face between TM helices 3, 5, 6, and 7.[18] For monoamine GPCRs and some peptide GPCRs, there are a number of conserved residues such as the key acidic residue (Asp113) on TM3 that binds a basic nitrogen in both agonists and antagonists. Most chemokine GPCRs possess an acidic residue on TM7 (Glu291) that performs the same anchoring role.[19] The conserved nature of this TM binding site helps explain why some ligands are capable of binding to both monoamine and peptide GPCRs.

Despite these commonalities, there must clearly be differences between typical monoamine and peptide binding sites that explain the apparently lower efficiency of ligands for the latter. Much evidence suggests that monoamine GPCR ligands bind tightly to a well-defined crevice deep within the conserved TM regions. The structure of the peptide GPCR binding sites is likely to be more open and diffuse with the ligands projecting out of the conserved TM region in the direction of the extracellular face. Significant binding to the extracellular domains, especially the second and third extracellular loops, may occur. The literature contains several examples of similar ligands containing a privileged substructure that bind to either monoamine or peptide GPCRs, but where the peptide GPCR ligand is larger.[20] This size difference presumably arises from a requirement for peptide GPCR ligands to span a more diffuse binding site that reaches closer to the extracellular surface.

Property Changes during Optimization. There was much lower variability in the property changes during optimization across the target family groups than there was in the properties of the starting or optimized compounds. Target families that significantly differed in terms of the median property values for the starting compounds often shared the same or similar values for the degree of change of those properties during optimization (Tables 1 and 3−7). This means that the principal determinant of the observed differences between the families is the property profile of the starting compounds rather than the changes that occurred during optimization. Therefore, clustering of the target family subsets on the basis of the properties of the starting compounds will give a similar result to that presented in Figures 3 and 4. The correlation across the families between the median MWs of the starting compounds and the optimized compounds was very high ($R^2 = 0.961$), whereas the correlation between the median changes in MW and the median MWs of the optimized compounds was much lower ($R^2 = 0.316$). This relative lack of variability among the target families may be related to the fact that many of the optimizations are concerned not only with enhancing potency, which tends to increase property values, but also with improving DMPK properties, which tends to constrain values.

Table 10. Changes in MW during Optimization Classified According to the MW of the Starting Compound

MW range for starting compd	number of entries	median change in MW	mean change in MW
	all families		
<200	69	76	97
200−300	312	45	62
300−400	565	36	48
400−500	464	25	31
>500	270	7	8
	monoamine GPCRs		
<300	115	30	50
300−450	169	18	25
>450	42	10	−4
	peptide GPCRs		
<300	22	68	77
300−450	131	57	73
>450	156	18	22

Nonetheless, there were some noteworthy differences between the target families in terms of the property changes during optimization. For some of the peptidergic families, substantial increases in some of the properties during optimization exacerbated the already high values for the starting compounds. The exceptionally low property values for the transporter optimized ligands were due to a coincidence of low values for the starting compounds and lower than average increases during optimization.

Increasing molecular size during optimization was observed for all of the families. The degree of this MW increase has previously been reported to be dependent upon the MW of the starting compounds, with starting compounds that have higher MW giving lower increases in MW.[21] Larger starting compounds may be more highly optimized already, at least in terms of potency, and therefore, MW is less subject to upward pressure. In addition, medicinal chemists will often exert downward pressure on MW when dealing with large starting compounds to achieve drug-like properties. The full SCOPE data set showed this same trend with median MW increases of 76 and 7 for starting compounds with MWs of <200 and >500, respectively (Table 10). Within the monoamine or peptide GPCR subsets, this trend was also apparent. However, this trend is overridden by the powerful underlying effect of the target family based differences in MW, as illustrated by the higher MW increase for the larger peptide GPCR starting compounds compared to that of the smaller monoamine GPCR and transporter ligands (Table 1). This again implies that peptide

GPCR ligands are inherently less efficient than their aminergic counterparts with a larger increase in size being required during optimization.

Addressing the Difficult Families. These overall findings based upon the medians should not be taken to imply that medicinal chemists working on the most difficult families such as peptide GPCRs will be unsuccessful in finding one suitable ligand (or two) that can be progressed to the clinic. This cause for optimism is supported by the lower quartile data (Tables 1 and 3−7). The lowest MW and cLogP quartiles for the peptide GPCR family are 434 and 3.8, indicating that a substantial proportion of the ligands fall within the drug-like RO5 range.

This analysis takes no account of the inevitable differences between targets within the same target family subset. Within the peptide GPCR subset, there are some targets that are unusually tractable and for which many drug-like ligands exist, such as the opioid receptors. For other targets within the most challenging families, it may be very difficult indeed to find ligands with RO5-compliant physicochemical properties. In these cases, the medicinal chemist needs to adopt a pragmatic approach to rule-based concepts of drug-likeness and try to identify those islands of acceptable oral activity that are scattered across the nondruglike non-RO5 space. After all, there are many ligands with high MWs that fall well outside the RO5 space but still display acceptable oral bioavailability, such as the cyclic peptide, cyclosporin.

Because peptide GPCR ligands appear to have a lower basal level of efficiency, medicinal chemists must aim for the maximum level of ligand efficiency that is possible for a given chemotype binding to a given target. If there are parts of a ligand that are not essential for the overall profile of the molecule, these should be identified and modified or deleted. In other words, diffuse binding sites need to be spanned with molecular skeletons that are as minimalistic as possible.

Most certainly, this analysis of drug-likeness does not mean that peptide GPCRs will always be more challenging drug targets than transporters, but from a purely physicochemical perspective, there is a higher likelihood that orally bioavailable ligands will be obtained for the latter. The drug discovery process is complex and many other considerations come into play, such as the need for sufficient in vitro and in vivo potency, selectivity, duration of action, an acceptable safety profile, synthetic feasibility, and patentability, to name but a few.

Data Interpretation and Caveats. There are a number of caveats that should be taken into account when considering the implications of these findings. The SCOPE database contains predominantly pre-clinical compounds from the 2000−2005 period; therefore, some of these findings may only be valid for those targets and target families that were most highly reported during this period. There will also be a considerable amount of variability in the SCOPE data due first to a degree of ambiguity in identifying the most optimized compound from some publications and second to differences in the extent to which the optimized compound has been optimized. This variability might affect the absolute values of the properties to some extent, but nonetheless, it is encouraging to see generally good agreement between the SCOPE average property values and those reported by other researchers for pre-clinical compounds. This suggests that the SCOPE data set, despite its limited size, is representative of a wider diversity of pre-clinical compounds. In addition, the changes in the properties during optimization for the full SCOPE set are consistent with previous work. It is likely that the effect of any variability would be neutral across the target families with all of the subsets being affected to a similar extent and therefore, overall, this should not significantly affect the rank order.

The large differences between the properties of oral drugs and the compounds in the SCOPE database is not surprising because it has been previously reported that as compounds pass through the various stages of clinical development their average MW tends to decrease, and the SCOPE database predominantly contains pre-clinical compounds.[4] Another consideration is that the targets and target families represented in the SCOPE database, which predominantly covers the 2000−2005 period, are likely to be different from those addressed by marketed drugs. Target families with high property values, such as peptide GPCRs and integrin receptors, are not well represented in current databases of marketed drugs.

Conclusions

A statistical analysis of a proprietary database of 1680 optimizations (SCOPE) showed that target families can be discriminated and clustered on the basis of the average physicochemical properties of their ligands. These differences between families influence the relative feasibility of obtaining oral drugs. For some families, many of the optimized ligands possessed properties that transgressed the limits of the drug-like RO5 space. For example, peptide GPCR ligands had a median MW of 510 and a cLogP of 5.0. There were deep rooted trends that prevailed, no matter how the data for the individual physicochemical properties were interpreted. The ligands for peptide GPCRs, integrin receptors, proteases, and transferases possessed high median property values, whereas the ligands for monoamine GPCRs, oxidases, and transporters, consistently, had low values. The median values for the optimized compounds for the transporters were so low that they often fulfilled the so-called rule-of-3 for fragment-based drug discovery.[22] Given that the ligands for different families could be categorized on the basis of their properties, the property ranges described here might be useful in the design of libraries targeted at particular families.

The properties of receptor agonists were generally more favorable than those for receptor antagonists and enzyme inhibitors. However, this effect was restricted to the agonists for monoamine GPCRs, opioid receptors, and ion channels, with no difference being found between the agonists and antagonists for non-opioid peptide GPCRs.

The degree of change of the properties during optimization was found to be fairly constant across the target families, meaning that the properties of the starting compounds were most critical in determining the properties of the optimized compounds.

There was a strong connection between the nature of the endogenous ligand and the complexity of the optimized synthetic ligands, with peptide-binding targets giving ligands with less favorable properties. In terms of affinity (p*K*i) per unit of MW or cLogP, the monoamine GPCR ligands bound with higher efficiency than the peptide GPCR ligands. These conclusions are certainly intuitive and consistent with the experience of many medicinal chemists whose aim was an oral drug for a peptide receptor. Despite the extra challenges presented by these difficult targets, experience has shown that ligands with good oral bioavailability can still be found, though the search for those rare compounds may be long and tortuous.

Acknowledgment. Key roles in the development of the SCOPE database were played by Colin Gray, who did the programming, and Corinne Kay, John Clark, and many other

members of the SCOPE team who populated the database with the optimizations. I also thank Zoran Rankovic and George Gettinby for many invaluable discussions relating to this article.

Supporting Information Available: Differences in the median property values and related *p* values for the target families, Pearson correlation coefficients for the six physicochemical properties, and the definitions of hydrogen bond donors and acceptors and rotatable bonds. This material is available free of charge via the Internet at http://pubs.acs.org.

References

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

(2) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem., Int. Ed.* **1999**, *38*, 3743−3747.

(3) 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.

(4) Wenlock, M. C.; Austin, R. P.; Barton, P.; Davis, A. M.; Leeson, P. D. A Comparison of physicochemical property profiles of development and marketed oral drugs. *J. Med. Chem.* **2003**, *46*, 1250−1256.

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

(6) 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.

(7) 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.

(8) Definitions of the methods used for calculating the numbers of hydrogen bonding groups and rotatable bonds are provided as Supporting Information.

(9) Blake, J. F. Examination of the computed molecular properties of compounds selected for clinical development. *BioTechniques* **2003**, *34*, S16−S20.

(10) 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.

(11) The Mann−Whitney point differences in the medians and the *p* values for the comparisons between target families are provided as Supporting Information.

(12) 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.

(13) There was a strong correlation (Pearson coefficient >0.8) between MW and RB and between PSA and HBA. The least correlated and most independent properties were cLogP and HBD. A table of Pearson coefficients of correlation for all 6 properties is provided in the Supporting Information.

(14) Black, J. Drugs from emasculated hormones: the principle of syntopic antagonism. *Science* **1989**, *245*, 486-493

(15) Beeley, N. R. A. Can peptides be mimicked? *Drug Discovery Today* **2000**, *5*, 354−363.

(16) Beaumont, K.; Schmid, E.; Smith, D. A. Oral delivery of G protein-coupled receptor modulators: An explanation for the observed class difference. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3658−3664.

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

(18) Klabunde, T.; Hessler, G. Drug design strategies for targeting G-protein-coupled receptors. *ChemBioChem* **2002**, *3*, 928−944.

(19) Mirzadegan, T.; Diehl, F.; Ebi, B.; Bhakta, S.; Polsky, I.; McCarley, D.; Mulkins, M.; Weatherhead , G.S.; Lapierre J. M.; Dankwardt J.; Morgans D.; Wilhelm R.; Jarnagin K. Identification of the binding site for a novel class of CCR2b chemokine receptor antagonists: binding to a common chemokine receptor motif within the helical bundle. *J Biol Chem.* **2000,** *275*, 25562−25571.

(20) Bondensgaard, K.; Ankersen, M.; Thogersen, H.; Hansen, B. S.; Wulff, B. S.; Bywater, R. P. Recognition of privileged structures by G-protein coupled receptors. *J. Med. Chem.* **2004**, *47*, 888−899.

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

(22) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'rule of three' for fragment-based lead discovery? *Drug Discovery Today* **2003**, *8*, 876−877.

JM0512185