JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

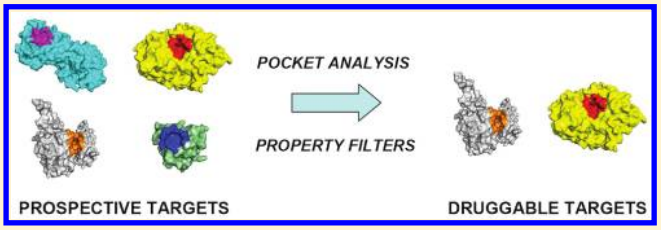
Development of a Rule-Based Method for the Assessment of Protein Druggability

Emanuele Perola,*[,†] Lee Herman,[†,§] and Jonathan Weiss[†]

[†]Vertex Pharmaceuticals, 130 Waverly Street, Cambridge, Massachusetts 02139, United States

Supporting Information

ABSTRACT: Target selection is a critical step in the majority of modern drug discovery programs. The viability of a drug target depends on two components: biological relevance and chemical tractability. The concept of druggability was introduced to describe the second component, and it is defined as the ability of a target to bind a drug-like molecule with a therapeutically useful level of affinity. To investigate the rules that govern druggability, we developed an algorithm to isolate and characterize the binding pockets of protein targets. Using this algorithm, we performed a comparative analysis between the relevant pockets of 60 targets of approved drugs and a diverse set of 440 ligand-binding pockets. As a result, we defined a preferred property space for druggable pockets based on five key properties (volume, depth, enclosure, percentage of charged residues and hydrophobicity), and we represented it with a set of simple rules. These rules may be applicable in the future to evaluate the chemical tractability of prospective targets.



INTRODUCTION

Based on a recent estimate, over 60% of the failures in drug discovery programs are due to poor choice of target.[1] This observation underscores the potential impact of reliable target evaluation methods on the notoriously high attrition rate in drug discovery. The viability of a protein as a drug target is determined by two components: biological relevance and chemical tractability. If modulation of the target's activity can lead to the desired biological effect, the target is biologically viable. However, it is also key that such modulation can be achieved using a drug-like molecule. When that is possible, the target is considered chemically tractable. The concept of druggability was introduced to address the second component and can be defined as the ability of a protein to bind a drug-like molecule with a therapeutically useful level of affinity.

A variety of methods and approaches to predict or estimate protein druggability have been reported in recent years. Cheng and colleagues at Pfizer developed a method to estimate the maximum achievable affinity for a given pocket based on its structure,[2] Hajduk and colleagues at Abbott derived a regression-based model to correlate pocket properties to experimental NMR screening rates,[3] Sheridan and colleagues at Merck derived an equation to estimate the "bindability" of a protein pocket based on three descriptors,[4] and similar functions to estimate druggability were developed by Halgren at Schrödinger[5] and by Schmidtke and Barril at the University of Barcelona.[6] Two additional approaches were recently published after the present study was completed: Krasowski and colleagues at the University of Dundee in collaboration with Schmitt at AstraZeneca used partial least-squares discriminant analysis to derive a druggability function based on 5 binding site descriptors;[7] Volkamer and Rarey at the University of Hamburg, in collaboration with Kuhn and colleagues at Merck Serono, developed an SVM-driven druggability model based on a series of global pocket descriptors, complemented with a local nearest neighbor model based on the distributions of pairwise distances between atom types of similar and different polarity within the pocket.[8]

In this study we took a slightly different approach that mirrors the type of analysis Lipinski and colleagues conducted on oral drugs to derive the well-known "rule of 5".[9] The goal was to derive a set of simple, intuitive and reliable rules defining the preferred property space for druggable pockets. While most of the methods reported in the literature were explicitly trained to recognize targets of oral drugs, we did not impose restrictions on the route of administration. In order to achieve our objective, we took the following steps: 1) identified a set of confirmed drug targets with available crystal structures; 2) compiled a representative subset of all ligand-bound proteins in the PDB; 3) developed an algorithm to isolate the relevant pocket and calculate its properties; 4) applied the algorithm to the proteins in the drug target and PDB sets; 5) performed a statistical analysis to identify the properties that provide the best discrimination between the two sets; and 6) derived preferred ranges for each of the selected properties as prospective guidelines to estimate the druggability of novel targets. In order to assess the potential impact of our method, we compared it with the druggability prediction tools included in Sitemap[5] and Fpocket,[6] the cavity detection and characterization programs developed at Schrödinger and at the University of Barcelona, respectively. Finally, we built a test set of novel targets of known structure

Received: December 20, 2011
Published: March 26, 2012

dx.doi.org/10.1021/ci200613b | J. Chem. Inf. Model. 2012, 52, 1027−1038

Table 1. Composition of the Drug Targets Set

target	first drug (approval date)	PDB code
plasminogen	Aminocaproic acid (1964)	1b2i
beta-lactamase	Clavulanic acid (1984)	1blh
thyroid hormone receptor	Liothyronine (1956)	1bsx
transpeptidase	Penicillin (1947)	1ceq
cyclophyllin A	Cyclosporin (1983)	1cwa
vitamin D nuclear receptor	Calcitriol (1978)	1db1
thrombin	Argatroban (2000)	1dwc
cyclooxygenase 1	Aspirin (1899)	1eqg
estrogen receptor	Diethylstilbestrol (1941)	1err
acetylcholinesterase	Tacrine (1993)	1eve
PPAR gamma	Troglitazone (1997)	1fm6
bacterial urease	Acetohydroxamic acid (1983)	1fwe
human neutrophile elastase	Sivelestat (2002)	1h1b
HIV-1 protease	Saquinavir (1995)	1hpv
HMG-CoA reductase	Lovastatin (1987)	1hwk
C-abl kinase	Imatinib (2001)	1iep
renal dipeptidase	Cilastatin (1985)	1itu
phenylalanine hydroxylase	Sapropterin (2007)	1j8u
isoleucyl-tRNA synthetase	Mupirocin (1987)	1jzs
DNA gyrase B	Novobiocin (1964)	1kzn
glutamate AMPA receptor 2	Cyclothiazide (1963)	1lbc
EGFR kinase	Gefitinib (2003)	1m17
glucocorticoid receptor	Prednisone (1955)	1m2z
rac1	Azathioprine (1968)	1mh1
glycinamide ribonucleotide formyltransferase	Pemetrexed (2004)	1njs
influenza neuraminidase	Zanamivir (1999)	1nnc
neutral endopeptidase	Racecadotril (1993)	1r1j
phosphodiesterase 3	Milrinone (1987)	1so2
4-hydroxyphenylpyruvate dehydrogenase	Nitisinone (2002)	1t47
histone deacetylase	Vorinostat (2006)	1t69
phosphodiesterase 5a	Sildenafil (1998)	1tbf
calcineurin/FKBP12	Tacrolimus (1994)	1tco
bRAF kinase	Sorafenib (2005)	1uwh
angiotensin converting enzyme	Captopril (1978)	1uzf
HIV-1 reverse transcriptase	Nevirapine (1996)	1vrt
dipeptidyl peptidase IV	Sitagliptin (2006)	1x70
retinoic acid receptor beta	Tamibarotene (2005)	1xap
bacterial RNA polymerase	Rifamycin (1960)	1ynn
mineralocorticoid receptor	Spironolactone (1982)	2aa5
androgen receptor	Bicalutamide (1995)	2axa
ROCK1 kinase	Fasudil (1995)	2esm
proteasome S20 unit	Bortezomib (2003)	2f16
troponin C	Levosimendan (2000)	2kfx
maltase glucoamylase	Acarbose (1995)	2qmj
beta2 adrenoreceptor	Salbutamol (1968)	2rh1
renin	Aliskiren (2007)	2v0z
factor Xa	Rivaroxaban (2008)	2w26
progesterone receptor	Mifepristone (1988)	2w8y
PPAR-alpha	Bezaflbrate (1977)	2znn
Na+/K+ATPase	Digoxin (1975)	3a3y
vEGFR2 kinase	Sunitinib (2006)	3c7q
adenosine A2A receptor	Theophylline(1976)	3eml
cKIT kinase	Pazopanib (2009)	3g0e
Src kinase	Dasatinib (2006)	3g5d
phosphodiesterase 4	Ibudilast(1990)	3gwt
RXR alpha	Bexarotene (1999)	3h0a
carbonic anhydrase II	Acetazolamide (1953)	3hs4
adenosine deaminase	Pentostatin (1991)	3iar
arachidonate 5-lipoxygenase	Zileuton (1996)	3o8y
cyclooxygenase 2	Celecoxib (1998)	4cox

associated with compounds in late stage clinical trials, and we used it both as a validation set and as an unbiased test set for comparison with the other methods.

METHODS

Selection of the Drug Target Set. The PDB, a number of drug-related databases (MDDR,[10] DrugBank,[11] Integrity,[12] etc.), and various other literature and online sources were mined to identify confirmed targets of approved drugs with one or more publicly available crystal structures. The resulting structures were analyzed, and 60 targets that accommodate a single ligand in the drug-binding pocket were included in the drug target set. Fifty-three of these targets (88.3%) are associated with at least one orally administered drug, while the remaining 7 are associated with drugs with other routes of administration (see breakdown in the Supporting Information). The selected targets include 40 enzymes and 20 proteins from other classes, including receptors, ion channels and structural proteins. One representative crystal structure for each target was selected based on the criteria discussed in the "Results and Discussion" section. The composition of the drug target set is reported in Table 1.

Selection of the Reference Set. The reference set was selected from the refined set of the 2008 version of the PDBbind database,[13] which contained 1401 structures of protein-small molecule complexes with resolution <2.5 Å and ligands of MW < 1000. A subset of 440 structures, encompassing one representative structure for each individual protein in the set, was designated as the reference set. The composition of the reference set is provided in the Supporting Information.

Selection of Drug Targets with Multiple Drug-Bound Structures. A list of all the ligand-bound crystal structures available for the 60 proteins in the drug target set was extracted from the PDB. Names and structures of the bound ligands were matched against a list of approved drugs derived from a number of drug databases. As a result, all the structures containing an approved drug bound to its target were identified. Nine targets were found with three or more available drug-bound structures, and the corresponding structures were extracted and included in the set. The composition of this set is reported in Table 2.

Analysis of Conformational Variability. For each of the nine targets in the section above, the selected structures were aligned using the "align_binding_sites" script provided by Schrodinger, LLC.[14] The script performs a pairwise superposition of multiple protein structures using the C-alpha atoms of the residues within 5 Å of the ligand in the reference structure. For each target, the representative structure originally included in the drug target set was used as the reference. The root-mean-square deviation of the coordinates of the binding site residues was then calculated between each structure and the reference structure using internally developed software. The same residues used for the alignment were included in the calculation, and all heavy atoms were considered.

Selection of the Phase III Target Set. The protein targets associated with small molecules currently undergoing phase III

dx.doi.org/10.1021/ci200613b | J. Chem. Inf. Model. 2012, 52, 1027−1038

Table 2. Crystal Structures of Drug Targets with 3 or More Drug-Bound Structures Available in the PDB

target	bound drug	PDB code
angiotensin converting enzyme	Lisinopril	1o86
angiotensin converting enzyme	Enalaprilat	1uze
angiotensin converting enzyme	Captopril	1uzf
acetylcholinesterase	Tacrine	1acj
acetylcholinesterase	Donepezil	1eve
acetylcholinesterase	Galantamine	1qti
carbonic anhydrase II	Brinzolamide	1a42
carbonic anhydrase II	Dorzolamide	1cil
carbonic anhydrase II	Acetazolamide	3hs4
cyclooxygenase 1	Ibuprofen	1eqg
cyclooxygenase 1	Flurbiprofen	1eqh
cyclooxygenase 1	Alclofenac	1ht8
cyclooxygenase 2	Diclofenac	1pxx
cyclooxygenase 2	Celecoxib	3ln1
cyclooxygenase 2	Naproxen	3nt1
cyclooxygenase 2	Flurbiprofen	3pgh
cyclooxygenase 2	Indomethacin	4cox
dipeptidyl peptidase IV	Sitagliptin	1x70
dipeptidyl peptidase IV	Linagliptin	2rgu
dipeptidyl peptidase IV	Saxagliptin	3bjm
HIV-1 protease	Amprenavir	1hpv
HIV-1 protease	Ritonavir	1hxw
HIV-1 protease	Indinavir	2bpx
HIV-1 protease	Lopinavir	2q5k
HIV-1 protease	Nelfinavir	3ekx
HIV-1 protease	Atazanavir	3eky
HIV-1 protease	Saquinavir	3oxc
HIV-1 protease	Darunavir	3qoz
HIV-1 reverse transcriptase	Efavirenz	1fk9
HIV-1 reverse transcriptase	Nevirapine	1vrt
HIV-1 reverse transcriptase	Etravirine	3mec
HIV-1 reverse transcriptase	Rilpivirine	3mee
phosphodiesterase 5a	Sildenafil	1tbf
phosphodiesterase 5a	Tadalafil	1xoz
phosphodiesterase 5a	Vardenafil	1xp0

clinical trials were extracted from the Integrity database.[12] Eleven targets were selected based on the following criteria: 1) target not associated with an already approved drug, 2) target structure available from the PDB, and 3) a single ligand bound in the relevant pocket. One structure per target was selected according to the same criteria used for the drug target set. Nine of the 11 targets are associated with orally administered drug candidates. The composition of the phase III target set is reported in Table 3.

Table 3. Composition of the Phase III Targets Set

target	drug candidate	PDB code
Cathepsin K	Odanacatib	1au3
c-MET kinase	Tivantinib	1r0p
fibroblast growth factor receptor 1	Brivanib alaninate	1agw
IGF-1R	OSI-906	1jqh
insulin receptor	OSI-906	1ir3
Jak2	Ruxolitinib	2b7a
Jak3	Tasocitinib	3pjc
methylguanine methyltransferase	O6-Benzylguanine	1eh8
Plk-1	ON-1910Na	3kb7
purine nucleoside phosphorylase	Forodesine	1pf7
Syk	Fostamatinib	3fqs

Isolation and Characterization of the Binding Pockets. The crystal structure of each protein−ligand complex was processed using the stepwise procedure described below. The outcomes of each step for a representative test case are illustrated in Figure 1.

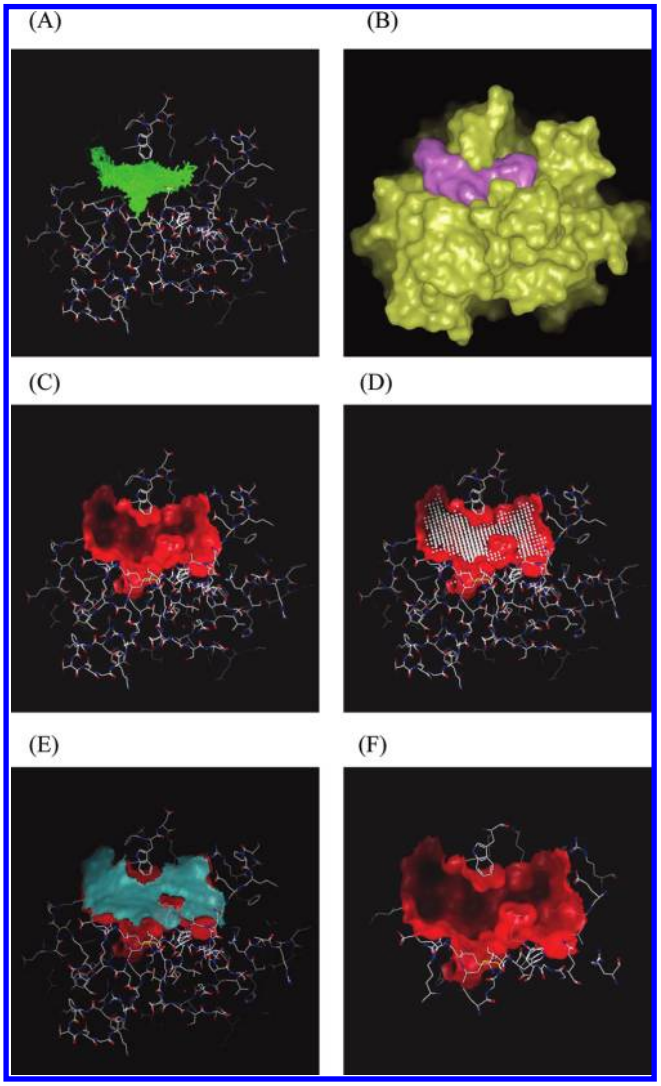


Figure 1. Graphical representation of the procedure applied to isolate the binding pockets. (A) 100 fragment-size compounds (green) are docked into the relevant pocket. (B) van der Waals surfaces are generated around the protein (yellow) and the ensemble of docked ligands (pink). (C) The pocket surface is selected (red). (D) The pocket is filled with site points (white). (E) A virtual lid is generated to cap the entrance of the pocket (cyan). (F) The pocket residues are selected.

Docking of Fragment Set. The structure downloaded from the PDB is processed as follows. The polypeptide component of the protein, any tightly bound waters (i.e., water molecules forming three or more hydrogen bonds to protein and ligand combined) and any active site metals are retained, while salts, counterions and buffer molecules are removed. The protonation state of the protein is assigned using the applyhtreat script provided by Schrödinger, LLC,[14] while the protonation state of the ligand is assigned using a script developed at Vertex that ionizes common functional groups to reproduce behavior at physiological pH. The hydrogen atoms in the complex are energy minimized with the Macromodel program[15] using the

dx.doi.org/10.1021/ci200613b | J. Chem. Inf. Model. 2012, 52, 1027−1038

OPLS2005 force field,[16] while the heavy atoms are kept fixed in their original crystallographic positions. The protein and the crystallographic ligand are then saved to separate files. Tightly bound waters and metals are included in the protein file. A diverse set of 100 fragment-size compounds (MW 100−350) is then docked into the pocket originally occupied by the X-ray ligand with the Glide program[17,18] in standard precision mode using default settings. One pose per compound is generated. The fragment set used in this study was selected from the Vertex fragment library using an internally developed program that selects a user-defined number of maximally dissimilar compounds from a collection. The purpose of this docking step is to fill the pocket evenly and then use the ensemble of docked compounds to guide the initial pocket definition step. In order to minimize the effect of outliers that may extend beyond the actual pocket, a grid with 2 Å spacing is generated around the docked compounds, grid cells that contain atoms from less than three compounds are identified, and the corresponding atoms are deleted.

Selection of Pocket Surface. After temporary removal of conserved water molecules and metal ions, a surface is generated around the protein and a separate surface is generated around the ensemble of docked ligands using a 1.05 Å radius probe. The resolution of the surface points is set to 0.5 Å. This and all subsequent surface generation tasks are performed using the Spicoli toolkit from OpenEye Scientific Software.[19] An initial set of protein surface points is selected and assigned to the pocket surface based on the following criteria: a) distance from the closest ligand surface point <5 Å and b) angle between the surface normal vector at the point under consideration and the surface normal vector originating from the ensemble ligands surface at the point of contact with the first vector >170 degrees. Additional steps are then taken to extend and refine the surface of the pocket: a) all initial points neighboring a surface point that is not part of the initial selection (edge points) are removed; this step was empirically determined to prevent the inclusion of neighboring pockets separated by a small ridge from the main pocket, b) surface points neighboring the originally selected points are evaluated. Normal vectors to the surface are generated for each new point under consideration and for its immediate neighbors from the initial selection. If the angle between the normal to the new point and the normal to one of its neighbors is less than 35 degrees the new point is added to the initial selection. New points continue to be added in successive iterations until the angle between the surface normal at a prospective new point and the surface normal at the closest point from the original selection exceeds 35 degrees. A number of additional iterations are performed to ensure that neighboring pockets are excluded from the main pocket and that all subpockets are fully included: 5 iterations of removal of pocket points neighboring nonpocket points, 15 iterations of addition of nonpocket points neighboring pocket points and 10 iterations of removal of pocket points neighboring nonpocket points. The number of connected sets of pocket surface points (sets representing continuous uninterrupted surfaces) is then determined, and the connected set containing the largest number of points is designated as the final pocket surface.

Generation of Pocket Site Points. In the following step the pocket is filled with site points using the following procedure. A grid with 0.7 Å spacing is generated around the pocket surface, and each grid point is evaluated. If a grid point is outside the protein and the distance to the closest pocket surface point is between 1 and 8 Å, a site point with 1 Å radius centered on that grid point is generated. The site points thus generated are retained if their degree of burial is higher than 55%. This value was empirically determined based on visual inspection of the results obtained with different thresholds. The degree of burial of each site point is determined as follows. A grid with 0.5 Å spacing is generated around the protein. A set of 72 equally spaced facet points are generated around the site point. An equal number of rays are projected from the site point, each connecting the site point with one of the 72 facet points. Each ray is extended in 0.25 Å increments until it encounters a cell containing a pocket surface point, up to a maximum distance of 50 Å from the corresponding site point. If less than 40 of the 72 rays hit the pocket surface (degree of burial <55%), the site point is deleted. To determine the final set of site points, an additional step is performed to ensure that only one group of site points is selected. To that end, a molecular surface with 0.05 Å resolution is generated around the site points using a 0.1 Å radius probe. If multiple connected surfaces are generated, only the site points corresponding to the surface with the largest area are retained.

Generation of Pocket Lid. For the purposes of determining the pocket depth, a virtual lid is generated to define the entrance of the pocket using the following procedure. A hybrid molecule composed of the protein atoms and the pocket site points is generated. This hybrid molecule is wrapped in a surface with 0.5 Å resolution generated using a 1.05 Å radius probe. The surface of the lid is defined by the surface points for which at least one site point is closer than the closest protein atom. If there are no lid points, the pocket is fully enclosed. If there are lid points, the algorithm checks whether there are multiple connected sets of lid points, which would indicate that the pocket has more than one entrance. If there are multiple lids, one is selected as the main entrance for the purpose of depth measurement based on the following procedure. For each lid, the lid surface point closest to the lid center of mass is found. The lid center point is selected as the site point nearest to the lid surface center point. The largest lid is selected by default, as long as the lid center point is not more than 76% buried. This value was empirically determined based on visual inspection of the results obtained with different thresholds. The degree of burial is computed with the same algorithm used for the selection of site points but using all the surface points instead of the pocket surface points only. In this case, if a smaller lid is at least 60% the size of the largest lid and its center point has a lower degree of burial, the smaller lid is selected instead.

Refinement of Pocket Surface. One extra step is performed at this stage to further refine the selection of pocket surface points. Pocket surface points that are more than 3 Å away from the closest site point are removed. Then, for each pocket surface point that neighbors a nonpocket surface point the distance to the closest site point is measured, and if the distance is over 1.9 Å the pocket surface point is removed. The process is continued until all edge points are within 1.9 Å of their closest site points.

Selection of Pocket Residues. The actual binding pocket is defined by the residues associated with pocket surface points. Each pocket surface point is associated with its closest protein atom and with the corresponding residue. If the backbone atoms of a given residue are associated with pocket surface points but the side chain atoms are not, the residue is converted to glycine. A surface is generated around the pocket residues

and the site points; the conserved waters and metal ions that fall within this surface and were removed at the initial stages of the process are added back to the pocket.

Computation of Pocket Properties. For each pocket in the drug target set and in the reference set we calculated the following properties: volume, depth, enclosure, surface area, percentage of charged residues, percentage of aromatic residues and average hydrophobicity. Aspartic acid, glutamic acid, lysine and arginine were treated as charged, and metal ions were included in the count of charged residues whenever present in the drug-binding pocket. The hydrophobicity was calculated as the average value of the Kyte-Doolittle hydropathy index[20] over all the pocket residues. The values of this index for each amino acid residue are listed in Table 4. The pocket volume was calculated as the volume enclosed by the surface generated around the pocket site points (Figure 2A). The pocket depth was computed as the distance between the closest pocket site point to the lid center and the furthest pocket surface point from any lid point (Figure 2B). For fully enclosed pockets (no lid) the depth was computed as the maximum distance between two pocket surface points. The degree of burial was computed as a function of the surface areas of pocket and lid (Figure 2B) using the following formula:

$$\text{Burial} = 2*(\text{pocket surface area}/(\text{pocket surface area} + \text{lid surface area})) - 1$$

Table 4. Kyte-Doolitle Hydropathy Indexes for the 20 Natural Amino Acids

RES	HYD	RES	HYD
ALA	1.8	LEU	3.8
ARG	−4.5	LYS	−3.9
ASN	−3.5	MET	1.9
ASP	−3.5	PHE	2.8
CYS	2.5	PRO	−1.6
GLN	−3.5	SER	−0.8
GLU	−3.5	THR	−0.7
GLY	−0.4	TRP	−0.9
HIS	−3.2	TYR	−1.3
ILE	4.5	VAL	4.2

Statistical Analysis. The property distributions for the pockets in the drug targets set and in the reference set were generated using Microsoft Excel. The correlations between different properties were analyzed using the same program. The distributions of the individual properties across the two sets were compared using the Kolmogorov−Smirnov (K−S) test provided in the R software package, version 2.13.0.[21] For each property, the null hypothesis to be evaluated was that there is no difference between the distribution in the drug targets set and that in the reference set. The preferred property ranges for druggable pockets ("druggability rules") were calculated using an internally developed program. The program performs a systematic analysis of the property distributions in the two sets to determine the set of property ranges that covers 90% or more of the drug targets set and maximizes the difference in coverage relative to the reference set. Performance and stability of the resulting rules were evaluated using the following 5-fold cross-validation procedure. The drug targets set and the diverse set were combined, and the resulting set was randomly partitioned into 5 subsets of equal size, with the constraint that each contain 20% of the drug targets set and 20% of the reference set. Four of these subsets were combined and used as the training set, while the remaining subset was used as the validation set. The process was repeated 5 times; each time a different subset was selected as the validation set and the remainder of the data was used for training. The preferred property ranges were computed for each training set as described above and evaluated on the corresponding test sets. A randomization test was also conducted to determine whether the method we employed could generate an equally discriminating model between random subsets of the overall distribution. The drug targets and the diverse set were combined, and a random subset of 60 targets was extracted and designated as the positive set. The remaining 440 targets were designated as the reference set. The partition process was repeated five times, with different random subsets selected each time using Microsoft Excel. In each case the property ranges covering at least 90% of the positive set while maximizing the separation from the reference set were computed with the above-described method. The degree of separation thus obtained was then compared with the separation between the

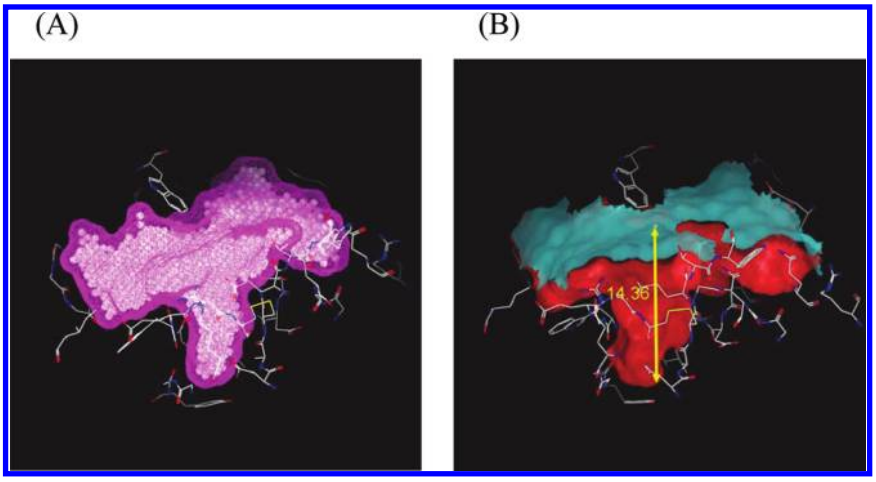


Figure 2. Graphical representation of the procedures applied to compute the 3D properties of the pocket. (A) Isolated pocket with corresponding site points. The pocket volume is computed as the volume enclosed by the surface generated around the site points (site points: white spheres; surface: magenta-colored mesh). (B) Isolated pocket with surface (red) and virtual lid (cyan). The pocket depth is computed as the distance between the lid center and the surface point farthest from any lid point (yellow line). The degree of enclosure is computed as a function of the surface areas of pocket and lid.

dx.doi.org/10.1021/ci200613b | J. Chem. Inf. Model. 2012, 52, 1027−1038

drug targets and reference sets provided by the druggability rules.

Sitemap Calculations. The Sitemap program (version 2.4) was obtained from Schrödinger, Inc. The same structures prepared for the docking step were processed using the command line version of the program. The crystallographic ligands were used to define the pockets of interest, and the calculations were performed with default settings. A druggability score was computed for each pocket.

Fpocket Calculations. The Fpocket software suite (version 2.0) was downloaded from http://fpocket.sourceforge.net. The calculations were conducted on the original PDB files using the dpocket program, which calculates pocket descriptors and druggability scores. The relevant pockets were specified using the 3-letter code of the crystallographic ligands, and default settings were applied.

RESULTS AND DISCUSSION

The factors affecting the druggability of protein pockets have been extensively studied in the past few years. As a result, a number of scoring functions for druggability prediction have been derived, most consisting of a linear combination of three or more parameters. The goal of this study was to enable a physically interpretable assessment of the druggability of a pocket by identifying the key properties that determine it and deriving a preferred range for each property. To that end, we developed a method to isolate and characterize protein pockets and used it to compare the properties of a set of proven druggable pockets (i.e., pockets associated with at least one approved drug) with those of a diverse set of ligand-binding pockets from the PDB.

Criteria for the Selection of the Data Sets. One critical aspect in this kind of study is the choice of the data that is used in the analysis. Given the goal of the study, the ideal choice in this case would have been to compare a set of validated drug targets with a set of proven undruggable targets. However, while validated drug targets can be identified based on their association with approved drugs, there is no reliable metric to conclusively prove lack of druggability. In two of the recent studies,[2,3] specific targets were classified as undruggable based on one of two criteria: 1) extensive drug discovery efforts were directed at these targets over the years with no success or 2) crystal structures of these targets were available but no high affinity drug-like ligands were reported for them in the literature. The set of targets meeting the first criterion is relatively small,[2] and their classification remains uncertain. A good example is HIV-1 integrase, classified as undruggable in the Pfizer study.[2] Raltegravir, an oral drug targeting this enzyme, was approved by the FDA the same year the paper was published. In the Abbott study,[3] 37 targets are classified as "low-druggability" based on the second criterion. However, the authors explicitly state that they could not differentiate between those targets for which a search for drug leads was attempted and failed and those for which such a screen was never performed. Given these uncertainties, we decided to make no assumptions on the lack of druggability of specific targets. Our broader assumption was that a large fraction of the potential protein targets are not druggable, and a similar proportion should be observed in a broad and diverse set of proteins representative of target space. While there is no way to conclusively demonstrate it, the success rates reported in the literature for high throughput screens against a broad range of targets strongly support this assumption. A survey published in 2006 analyzed the outcome of the screens conducted in 47 different HTS laboratories in 2004 and showed that viable leads were identified for about 50% of the targets.[22] In 2008 Novartis reported that 57% of their internal screens produced a workable lead series.[23] In a similar study, AstraZeneca analyzed 36 internal high throughput screens performed between 2001 and 2008 across a broad range of targets and reported a 47% success rate in identifying hits of sufficient quality to initiate a hit-to-lead program.[24] Finally, GlaxoSmithKline analyzed the results of 70 internal high-throughput screens against a variety of antibacterial targets and reported that only 16 screens gave rise to hits and only 5 of them resulted in viable lead series.[25] Admittedly, HTS is not the only available approach to identify viable hits, and there are certainly cases where other approaches can succeed when HTS fails. On average, however, the ability of HTS to identify viable hits for a given target from a large and diverse collection of compounds is a good indicator of the druggability of that target. Considering the above data and the high attrition rates in the path from lead series to approved drugs, it is safe to assume that the fraction of targets of pharmaceutical interest that are actually druggable is unlikely to exceed 50%. Based on these considerations, we decided to compare a set of validated drug targets with a diverse set of proteins representing the entirety of the target space, and we tried to derive a set of rules that would correctly identify the former as a significantly smaller subset of the latter. The PDBbind database was chosen as a source for the proteins in the diverse set for a number of reasons: the size of the database guarantees a high degree of diversity, each entry is a protein−ligand binary complex, the structures come individually preprocessed, and the ligand-binding pockets have already been identified and can be directly used in the analysis. The second question we needed to address was the following: which structure(s) should be chosen to represent each target when multiple crystal structures are available. In principle, there were two possibilities: 1) including all the available structures and computing the average properties for the relevant pocket; 2) selecting one representative structure for each target and using the conformation from that structure to calculate the properties of the relevant pocket. The first option presented a number of problems. First, the average properties would likely correspond to a nonexisting conformational state of the pocket. Second, different targets would be treated differently depending on the number and conformational diversity of the available structures. Third, the druggability of a pocket depends on the properties of its most druggable conformation rather than on the average properties of all the accessible conformations. For these reasons we decided against this option. The second option required a sensible criterion to select the representative structure for each target. Ideally, one should choose the most druggable conformation of the relevant pocket among all the available structures. However, there is no obvious way to rank the druggability of different conformations a priori. While the most druggable conformation of a pocket cannot be identified with certainty, every conformation that accommodates an approved drug is by definition a druggable conformation. Ultimately, our goal was to compile a set of proven druggable pockets. We therefore selected for each target a structure in complex with one of the associated drugs whenever one was available. When no structure of the target in complex with an approved drug was available, we visually inspected the available structures and selected the structure bound to the most similar ligand to one of the approved drugs for that target. Overall, 41 of the 60

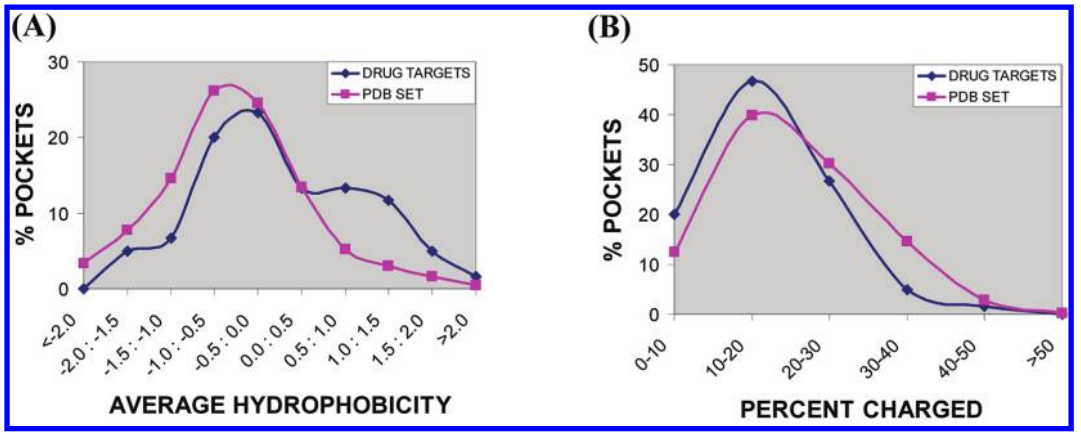


Figure 3. (A) Distributions of the average hydrophobicities for the pockets of the drug targets and reference sets. (B) Distributions of the percentage of charged residues for the two sets. The distributions are binned and have been represented as superimposed curves to best visualize the differences. The positions of the bins are defined by the square markers.

selected structures in the drug target set contain an approved drug in the relevant pocket, while most of the remaining 19 contain a ligand that shares the same or a similar scaffold with an approved drug for the same target. For one of the 60 targets (arachidonate-5-lipooxygenase) we included the apo structure, as it was the only one available, and we characterized the known active site which is also the binding pocket for the corresponding drug. For the diverse set from the PDBbind database, one structure per target was randomly selected.

Analysis of the Property Distributions. Each pocket in the drug target set and in the reference set was isolated and characterized as described in the Methods section. The properties we calculated can be divided into 1D properties (solely dependent on the composition of the pocket) and 3D properties (dependent on the size and shape of the pocket). For each property, the distribution in the drug targets set was compared with the distribution in the reference set to determine whether that property could be used to differentiate between the two sets. The comparison was based on the p-value obtained from the K−S test and on the magnitude of the differences observed at the margins of the two distributions. Our expectation was that the distributions of the key properties for druggable pockets would be either shifted or simply narrower relative to the corresponding distributions for the overall population. The objective was to identify an allowed or preferred range for each property, i.e. a range that would encompass the vast majority of the druggable pockets while excluding a significantly larger fraction of the overall population. A property was considered for inclusion in the model building step if the range of values covering 95% of the drug targets set encompassed less than 90% of the reference set, thus excluding over 10% of the reference set versus 5% of the drug targets set. Once a set of properties meeting these criteria was identified, pairwise correlations between these properties were computed to assess dependency and eliminate redundancy. Based on this analysis, five of the seven calculated properties were selected to represent the preferred property space for druggable pockets: hydrophobicity, percentage of charged residues, volume, depth and degree of enclosure. Surface area was excluded due to its high correlation with volume, while the degree of aromaticity did not provide sufficient discrimination between the two sets to be selected.

The distributions of the selected 1D properties are illustrated in Figure 3. As shown in Figure 3A, the pockets of drug targets tend to be more hydrophobic than average. The median hydrophobicities are −0.13 for the drug target set and −0.55 for the reference set, and the corresponding p-value obtained from the K−S test is 0.001. Highly hydrophobic pockets (average hydropathy index >1) are considerably more common in the drug target set than in the reference set, while for highly polar pockets (average hydropathy index < −1) the situation is reversed. The charge distributions for the two sets, illustrated in Figure 3B, are more similar to each other, but appreciable differences can be observed at the margins. The median percentages of charged residues in the binding pockets are 16.8 for the drug target set and 20.0 for the reference set, and the corresponding p-value is 0.009. Highly charged pockets are much less common in the drug target set than in the reference set, while for pockets with few or no charged residues the ratio is again reversed. The distributions of the selected 3D properties are illustrated in Figure 4. The distributions of the pocket volumes are very similar for the drug target set and the reference set, while more significant differences between the two sets can be seen for the distributions of depth and enclosure. The median volumes are 1209 Å^3 for the drug target set and 1189 Å^3 for the reference set, while the median depths are 15.8 Å and 14.5 Å, respectively, and the median enclosures are 0.56 and 0.49, respectively. The corresponding p-values obtained from the K−S test are 0.639, 0.048, and 0.061 for volume, depth, and enclosure, respectively. For volume and depth appreciable differences can be observed at the lower ends of the distributions, while for enclosure the differences are visually apparent at both ends. A close look at the lower ends of the three plots shows that small volumes, low depths and low levels of enclosure are significantly less common in the drug targets set than in the reference set. While the K−S test does not support the choice of volume as a differentiating property, we felt that an adequate description of property space should include at least one descriptor for size. Given that the ultimate goal is to define a preferred range for each property for druggable pockets, a possible lack of differentiating ability would not invalidate the corresponding rule. Moreover, a close look at the numbers shows that only 1 of 60 binding pockets in the drug targets set has a volume <500 Å^3, while 33 of 440 binding pockets in the reference fall within the same range. It is

dx.doi.org/10.1021/ci200613b | J. Chem. Inf. Model. 2012, 52, 1027−1038

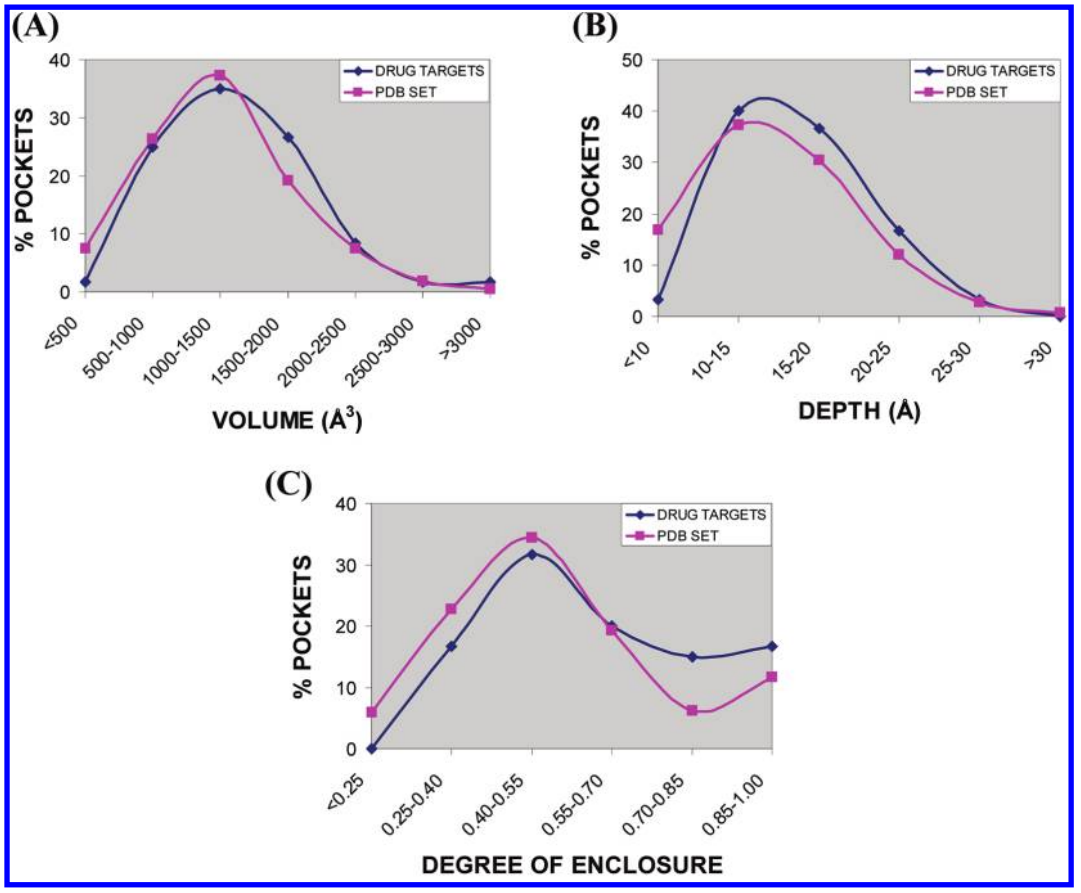


Figure 4. Distributions of the 3D properties for the pockets of the drug targets and reference sets: (A) volume, (B) depth, and (C) enclosure. The distributions are binned and have been represented as superimposed curves to best visualize the differences. The positions of the bins are defined by the square markers.

plausible that this difference reflects actual size requirements for druggable pockets and that the lack of significance suggested by the K–S test be due to the limitations of sample size and possibly to some intrinsic limitation in the ability of this test to compare the tails of distributions.[26]

The differences observed at the margins of each distribution for these five key properties are further illustrated in Table 5.

Table 5. Property Ranges Encompassing 95% of the Drug Targets Set and Percentage of Binding Pocket in the Reference Set Falling within the Same Ranges

property	range	% drug targets	% PDB set
volume (Å^3)	≥593	95	89.1
depth (Å)	≥10.6	95	81.1
enclosure	≥0.31	95	87.3
hydrophobicity	≥−1.24	95	83.6
percent charged	≤30.3	95	82.7

For each property the table reports the range of values that covers 95% of the drug targets set, expressed with a single threshold, and the percentage of the reference set that falls within the same range. The differences between the two percentages observed for the five properties suggested that a preferred property space for druggable pockets could be defined by using a set of thresholds and that an appropriate choice of thresholds could lead to a substantial discrimination between the drug target set and the reference set.

Emerging Guidelines. A systematic analysis of the property distributions was conducted to identify a set of thresholds that would cover at least 90% of the drug targets set while maximizing the separation relative to the diverse set. This analysis showed that 55 of the 60 binding pockets in the drug target set (91.7%) satisfy the following rules:

Volume ≥ 500 Å^3

Depth ≥ 10.4 Å

Enclosure ≥ 0.28

Percentage of charged residues ≤ 26.3

Hydrophobicity ≥ −1.12

Conversely, only 57% of the binding pockets in the diverse set satisfy the same rules. In order to validate this result, a 5-fold cross-validation was performed as described in the Methods section. As shown in Table 6, the thresholds calculated for the preferred property ranges were very stable across the 5 runs and provided consistent separation between drug targets and diverse targets in the corresponding training sets. The average values for the property thresholds and the fractions of targets falling within those thresholds were consistent with those obtained from the entirety of the data. The performance of the 5 models on the corresponding test sets was slightly more variable, but the separation provided between drug targets and diverse targets remained substantial in all cases. To further validate our model, we conducted a randomization test to

Table 6. Computation of Preferred Property Ranges for Druggable Pockets: 5-Fold Cross-Validation Results

	VOL	DPT	ENC	%CHG	HYD	%DT_train[a]	%PDB_train[b]	%DT_test[c]	%PDB_test[d]
Model1	≥502	≥10.4	≥0.29	≤26.4	≥−1.12	91.7	56.8	91.7	56.8
Model2	≥501	≥10.4	≥0.29	≤26.3	≥−1.12	91.7	56.2	91.7	59.1
Model3	≥501	≥10.4	≥0.29	≤26.2	≥−1.12	93.8	56.2	83.3	59.1
Model4	≥572	≥10.4	≥0.29	≤26.3	≥−1.12	91.7	55.7	83.3	53.4
Model5	≥501	≥9.3	≥0.30	≤26.4	≥−0.94	91.7	54.8	75	55.7
mean	**515**	**10.2**	**0.29**	**26.3**	**−1.08**	**92.1**	**55.9**	**85**	**56.8**
STD[e]	**32**	**0.5**	**0.004**	**0.08**	**0.08**	**0.9**	**0.7**	**7**	**2.4**

[a]Percentage of validated drug targets in the training set that fall within the calculated thresholds. [b]Percentage of diverse targets (targets selected from the reference set) in the training set that fall within the calculated thresholds. [c]Percentage of drug targets in the test set that fall within the calculated thresholds. [d]Percentage of diverse targets in the test set that fall within the calculated thresholds. [e]Standard deviation of the mean.

determine whether our method could artificially generate an equally discriminating set of rules between two random subsets of the overall distribution. The separation between random subsets obtained in each of 5 independent runs averaged 10.0 ± 6.7%, which is well below the separation our model provides between the actual drug targets and reference sets. These results confirm that the druggability rules described above are not an artifact of the computational approach used to derive them.

The significance of the separation provided by these rules is further illustrated by the following considerations:

1) The drug targets set and the reference set are not orthogonal: the reference set represents the entire spectrum of ligand-binding proteins, which includes confirmed drug targets, other druggable targets, and undruggable targets. These rules provide a separation between a subset and the entire population.
2) About 90% of oral drugs satisfy all the Lipinski rules, while about 69% of commercially available screening compounds satisfied the same rules at the time of their original publication (unpublished data). The number has increased since then to about 72%[27] as the Lipinski rules started to influence the choices of chemical vendors. The rules derived here are more discriminating for protein pockets than the Lipinski rules were for small molecules.
3) The reference set is derived from the PDBbind database, which only contains proteins bound to small molecules. It is therefore inherently biased toward targets that have the ability to bind small molecules, while it does not include other potentially more challenging targets for which a ligand-bound structure is not available. If the reference set included representatives of such targets, the separation provided by these rules would likely be even larger.

To further understand the significance of these rules, we analyzed the five violators in the drug targets set to identify any common patterns. Interestingly, the targets that violate the rules tend to be associated with drugs that are also in a potentially unfavorable region of property space. Figure 5 illustrates the structures of the drugs associated with the five targets that lie outside of the preferred property ranges. These drugs include two very large molecules that violate multiple Lipinski rules, a prodrug, a molecule with three ionizable centers and a drug that is unusually small. This observation suggests that targets that violate the rules may still be viable but will likely require drugs that are challenging to develop.

Impact of Structure Selection Criteria. To further validate our choice to represent each drug target with a single crystal structure, we identified 9 targets in our data set for

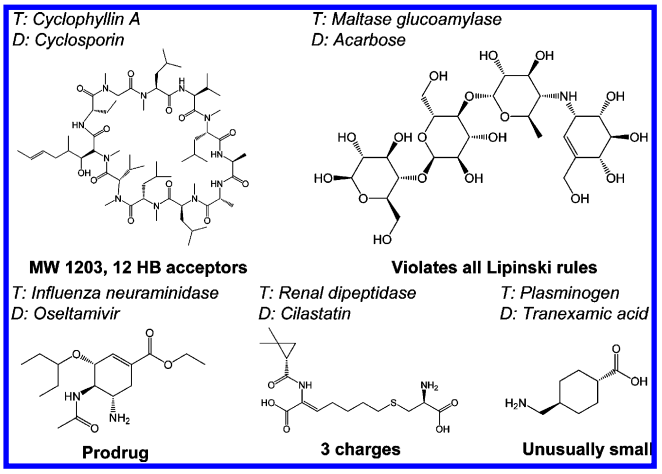


Figure 5. Structures of the drugs associated with the 5 targets in the drug target set that violate the druggability rules; *T* = target; *D* = drug.

which multiple crystal structures in complex with approved drugs are available and analyzed the conformational variability of the corresponding drug-binding pockets. The 35 structures available for these targets are listed in Table 2. For each target, the available structures were aligned using the structure included in the original data set as the reference. The alternative structures were then compared to the originally selected structure based on the root-mean-square deviation of the coordinates of the residues within the drug-binding pocket. Each of the alternative binding site conformations was within 1 Å of the initially selected conformation for all 9 targets with the only exception of one HIV-1 reverse transcriptase structure, which was 1.39 Å from the corresponding reference conformation. Visual inspection confirmed that the conformational variations in the binding site across the different drug-bound structures were minimal for all 9 targets, supporting the choice of a single representative structure and suggesting that the choice of alternative drug-bound structures would have been inconsequential to the results of the analysis. To further support this assessment, we applied our pocket detection and evaluation method to all the crystal structures in this set. The results showed minimal fluctuations of the key parameters across different drug-bound conformations for all 9 targets with very few significant deviations. The calculated values for the 5 parameters of choice were all within the preferred ranges defined in the previous section for 31 of the 35 structures, while marginal violations were observed for a single parameter in the remaining 4 structures. The calculated parameters for the structures in this set are included in the Supporting Information. This analysis supports the validity of our approach

dx.doi.org/10.1021/ci200613b | J. Chem. Inf. Model. 2012, 52, 1027−1038

and confirms that the guidelines emerging from this study are substantially independent of the specific drug-bound structures used to derive them.

Comparison with Sitemap and Fpocket. To assess the potential of our method relative to the current state-of-the-art approaches, we evaluated the ability of the available druggability prediction tools to differentiate between the drug targets set and the reference set. Two of the recently described methods are implemented in currently available software: Sitemap, developed at Schrödinger, Inc.,[5] and Fpocket, developed at the University of Barcelona.[6] Both of these tools identify and characterize protein pockets and compute a druggability score for each pocket as a function of its properties. As most reported druggability prediction methods, Sitemap and Fpocket were both trained to recognize targets of oral drugs, while we did not impose restrictions on the route of administration. In order to account for this difference, we extended this comparative assessment to the amended "oral only" drug targets set, which only includes the 53 drug targets in our training set that are associated with at least one oral drug.

Each structure in our sets was processed with both tools using default settings, and druggability scores were computed for the corresponding ligand-binding pockets. In some cases Fpocket identified multiple cavities in the ligand-binding pocket and calculated a score for each cavity. In those cases, we computed a global score for the ligand-binding pocket as the weighted average of the scores of its subpockets, each subpocket weighing proportionally to its volume. The cumulative distributions of the scores obtained with Sitemap and Fpocket for the drug targets set and the reference set are reported in Figure 6. In both cases, higher scores correspond to higher predicted druggability. As shown in Figure 6A, a threshold of 1.0 for the druggability score calculated with Sitemap yields a substantial separation between the two sets: 73.3% of the pockets in the drug target set have scores of 1.0 or higher, while only 41.4% of the pockets in the reference set satisfy the same criterion. However, if the score cutoff is reduced to increase the retrieval of drug targets (i.e., to reduce the false negative rate), the separation between the two sets becomes progressively smaller. For example, a score cutoff of 0.8 leads to the retrieval of 88.3% of the drug target set and 76.6% of the reference set. When the drug target set is replaced with the amended "oral only" subset, the results are almost identical: 73.6% of the pockets in the amended set have scores of 1.0 or higher, 90.6% have scores of 0.8 or higher. A similar trend is observed in the distributions of the scores calculated with Fpocket, as illustrated in Figure 6B. In this case a score threshold of 0.7 provides a substantial separation between the two sets, but the percentage of drug targets scoring above this threshold is low (50% vs 25.2% for the reference set). At lower cutoffs the retrieval of drug targets increases, but the separation between the two sets shrinks rapidly. In order to retrieve 86.7% of the drug target set the score cutoff needs to be lowered to 0.1, and 77.7% of the pockets in the reference set also score above the same cutoff. When the amended "oral only" subset is considered, the results are very similar: 52.8% of the pockets in the amended set have scores of 0.7 or higher, while 88.7% are retrieved when the score cutoff is lowered to 0.1. These results suggest that a Sitemap score of 1.0 or higher or an Fpocket score of 0.7 or higher would be required to classify a pocket as druggable with high confidence. However, in both cases (and particularly for Fpocket) the application of these cutoffs would result in a high false negative rate, as a large fraction of the confirmed druggable pockets score below these values. Conversely, the rules derived in this paper recall over 90% of the confirmed drug targets while excluding over 40% of the reference set, thus highlighting a potential advantage for the method described here. Considering that the training set used to derive our model was used as a test set for the other methods, we decided to extend our evaluation to include a truly external test set.

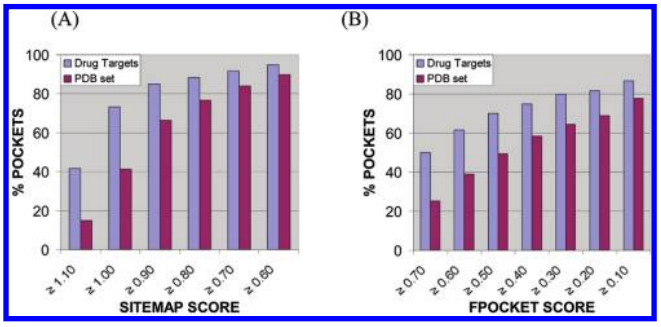


Figure 6. Distributions of the druggability scores calculated with Sitemap and Fpocket for the pockets of the drug targets and reference sets.

Evaluation of Novel Targets of Phase III Drug Candidates. In order to identify a naïve set to validate our model and conduct a more rigorous comparison with Sitemap and Fpocket, we searched for the targets of drug candidates that are currently in phase III clinical trials with the goal of compiling a test set of "quasi-validated" drug targets. As a result, we identified 11 novel targets with known crystal structures and a single ligand bound in the relevant pocket (Table 3). Nine of the 11 targets are associated with orally administered drug candidates. We then applied our method and the two external methods to each of these targets to characterize the relevant binding pockets. When our method was applied, 10 of the 11 targets (90.9%) satisfied all 5 druggability rules. The only target that violated at least one rule was Cathepsin K, a cysteine protease featuring a shallow and largely solvent exposed active site. Not surprisingly odanacatib, the drug candidate associated with this target, is a relatively large molecule (MW 525) that binds covalently to the enzyme (Figure 7). These results confirm that the model we derived classifies druggable pockets with a high degree of accuracy and that pockets outside the preferred property space are likely to require less conventional ligands with potential downstream liabilities. The druggability scores computed with Sitemap and Fpocket for this set are reported in Table 7. The distributions are consistent with those obtained on the drug targets set. When a score cutoff of 1 is

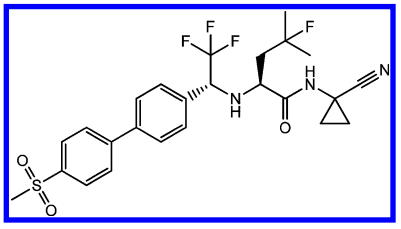


Figure 7. Structure of the drug candidate associated with cathepsin K, the only target in the phase III set that violates the druggability rules.

dx.doi.org/10.1021/ci200613b | J. Chem. Inf. Model. 2012, 52, 1027−1038

Table 7. Druggability Scores Computed with Sitemap and Fpocket for the Phase III Targets Set

target	Sitemap	Fpocket
Cathepsin K	0.82	0.30
c-MET kinase	1.08	0.40
fibroblast growth factor receptor 1	1.01	0.68
IGF-1R	0.94	0.60
insulin receptor	1.02	0.80
Jak2	1.10	0.72
Jak3	1.20	0.74
methylguanine methyltransferase	0.83	0.69
Plk-1	1.11	0.68
purine nucleoside phosphorylase	0.99	0.68
Syk	1.05	0.76

applied, Sitemap classifies 7 of the 11 targets as druggable (63.6%), while the other 4 targets score between 0.80 and 0.99. At the same cutoff, Sitemap recalls 6 of the 9 targets in the "oral only" subset (66.6%). When a cutoff of 0.7 is applied, Fpocket classifies 4 of the 11 targets as druggable (36.4%), while a cutoff of 0.5 recalls 9 out of 11 (81.8%). When the "oral only" subset is considered, Fpocket recalls 4 of the 9 targets at a cutoff of 0.7 (44.4%) and 7 of 9 at a cutoff of 0.5 (77.8%). These results confirm that, while Sitemap classifies druggable pockets with higher accuracy than Fpocket, a relatively large percentage of actual or likely drug targets cannot be classified as druggable with high confidence using either of these methods. This unbiased small scale evaluation is consistent with the results we obtained on the larger sets, suggesting that our method may indeed provide improved accuracy relative to the currently available software for druggability prediction. Given the limited size and diversity of the validation set, additional studies will be required for a conclusive assessment of the relative performance of these methods.

CONCLUSIONS

The assessment of protein druggability has been a hot topic in the pharmaceutical community for the past several years. The ability to triage new targets based on solid and validated criteria has the potential to reduce the high attrition rate that continues to plague drug discovery programs. In this study we strived to develop a method that was highly accurate in the isolation and characterization of the relevant binding pockets and that produced easily interpretable results. We then applied our method to the structures of a set of validated drug targets to identify the preferred region of property space for druggable pockets. As a result, we derived a set of rules that can be applied to the evaluation of prospective targets whenever the 3D structure and the location of the relevant pocket are known. While a conclusive assessment of the relative accuracy of different methods will require more testing, the results presented above show that our method fares well when compared to the currently available druggability prediction software. Moreover, the metric we derived is straightforward, intuitive and physically interpretable. As shown by the wide popularity of the Lipinski rule of 5, the impact of a computational method intended for broad use is heavily dependent on its simplicity, and that is one of the aspects we wanted to maximize in this study. We also believe that having an immediate read of the kinds of challenges a target may present can help make better decisions when the pocket of interest is close to the edge of the preferred property space. Combined with an accurate assessment of the related biology, this method has the potential to improve the target selection process and contribute to an increased success rate in future drug discovery programs. In order to facilitate the application of this method, we included the related software in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

Source code of the software developed and used to characterize the binding pockets. Tables containing composition and calculated parameters for the structures in the drug targets, PDBbind, phase III, and multiple conformations sets. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: (617) 444-6646. E-mail: emanuele_perola@vrtx.com.

Present Address

§Lee Herman Consulting LLC, 11 Penobscot Rd., Natick, MA 01760.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. W. Patrick Walters for valuable advice on key aspects of this study, Dr. Brian B. Goldman for help with software development, Dr. Brian McClain for the script to calculate the root-mean-square deviation between binding site residues of superimposed structures, Dr. Paul S. Charifson for careful reading and valuable feedback on the manuscript and Dr. Anthony Nicholls for valuable suggestions on the statistical validation of the method.

ABBREVIATIONS USED:

PDB, Protein Data Bank; K–S test, Kolmogorov–Smirnov test

REFERENCES

(1) Brown, D.; Superti-Furga, G. Rediscovering the sweet spot in drug discovery. *Drug Discovery Today* **2003**, *8*, 1067–1077.

(2) Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S. Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* **2007**, *25*, 71–75.

(3) Hajduk, P. J.; Huth, J. R.; Fesik, S. W. Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* **2005**, *48*, 2518–2525.

(4) Sheridan, R. P.; Maiorov, V. N.; Holloway, M. K.; Cornell, W. D.; Gao, Y. D. Drug-like density: a method of quantifying the "bindability" of a protein target based on a very large set of pockets and drug-like ligands from the Protein Data Bank. *J. Chem. Inf. Model.* **2010**, *50*, 2029–2040.

(5) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389.

(6) Schmidtke, P.; Barril, X. Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *J. Med. Chem.* **2010**, *53*, 5858–5867.

(7) Krasowski, A.; Muthas, D.; Sarkar, A.; Schmitt, S.; Brenk, R. DrugPred: A Structure-Based Approach To Predict Protein Druggability Developed Using an Extensive Nonredundant Data Set. *J. Chem. Inf. Model.* **2011**, *51*, 2829–2842.

(8) Volkamer, A.; Kuhn, D.; Grombacher, T.; Rippmann, F.; Rarey, M. Combining Global and Local Measures for Structure-Based Druggability Predictions. *J. Chem. Inf. Model.* **2012**, *52*, 360−372.
(9) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.
(10) *MACCS-II Drug Data Report 2011*; Accelrys: San Diego, CA, 2011.
(11) Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.; Djoumbou, Y.; Eisner, R.; Guo, A. C.; Wishart, D. S. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res.* **2011**, *39*, D1035−1041.
(12) *Thomson Reuters Integrity*; Thomson Reuters: New York, NY, 2011.
(13) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **2004**, *47*, 2977−2980.
(14) Schrodinger, LLC.: New York, NY, 2011.
(15) *Macromodel*, 9.9; Schrodinger L.L.C.: New York, NY, 2010.
(16) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(17) *Glide*, 5.6; Schrodinger L.L.C.: New York, NY, 2010.
(18) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739−1749.
(19) *OpenEye Spicoli toolkit*; OpenEye Scientific Software: Santa Fe, NM, 2010.
(20) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105−132.
(21) The R Project for Statistical Computing. www.r-project.org.
(22) Fox, S.; Farr-Jones, S.; Sopchak, L.; Boggs, A.; Nicely, H. W.; Khoury, R.; Biros, M. High-throughput screening: update on practices and success. *J. Biomol. Screen.* **2006**, *11*, 864−869.
(23) Bender, A.; Bojanic, D.; Davies, J. W.; Crisman, T. J.; Mikhailov, D.; Scheiber, J.; Jenkins, J. L.; Deng, Z.; Hill, W. A.; Popov, M.; Jacoby, E.; Glick, M. Which aspects of HTS are empirically correlated with downstream success? *Curr. Opin. Drug Discovery Dev.* **2008**, *11*, 327−337.
(24) Edfeldt, F. N.; Folmer, R. H.; Breeze, A. L. Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discovery Today* **2011**, *16*, 284−287.
(25) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discovery* **2007**, *6*, 29−40.
(26) Mason, D. M.; Schuenemeyer, J. H. A modified Kolmogorov-Smirnov test sensitive to tail alternatives. *The Annals of Statistics* **1983**, *11*, 933−946.
(27) Chuprina, A.; Lukin, O.; Demoiseaux, R.; Buzko, A.; Shivanyuk, A. Drug- and lead-likeness, target class, and molecular diversity analysis of 7.9 million commercially available organic compounds provided by 29 suppliers. *J. Chem. Inf. Model.* **2010**, *50*, 470−479.