Ligand-Based Approach to In Silico Pharmacology: Nuclear Receptor Profiling

Jordi Mestres,*,† Lidia Martín-Couce,† Elisabet Gregori-Puigjané,† Montserrat Cases,† and Scott Boyer‡

Chemogenomics Laboratory, Research Unit on Biomedical Informatics, Institut Municipal d'Investigació Mèdica and Universitat Pompeu Fabra, Dr Aiguader 80, 08003 Barcelona, Catalonia, Spain, and Safety Assessment, AstraZeneca R&D, 43183 Mölndal, Sweden

Received July 18, 2006

Bioactive ligands are a valuable and increasingly accessible source of information about protein targets. On the basis of this statement, a list of 25 nuclear receptors was described by a series of bioactive ligands extracted directly from bibliographical sources, stored properly in an annotated chemical library, and mathematically represented using the recently reported SHED molecular descriptors. Analysis of this ligand information allowed for derivation of a threshold of nuclear receptor concern. If the similarity of one molecule to any of the molecules annotated to one particular nuclear receptor is below that threshold, the molecule receives an alert on the probability of having affinity below 10 μM for that nuclear receptor. On this basis, a linkage map was constructed that reveals the interaction network of nuclear receptors from the perspective of their active ligands. This ligand-based approach to nuclear receptor profiling was subsequently applied to four external chemical libraries of 10 000 molecules targeted to proteases, kinases, ion channels, and G protein-coupled receptors. The percentage of each library that returned an alert on at least one nuclear receptor was reasonably low and varied between 4.4 and 9.7%. In addition, ligand-based nuclear receptor profiling of a set of 2944 drugs provided an alert for 153 drugs. For some of them, namely, acitretin, telmisartan, phenyltoloxamine, tazarotene, and flumazenil, bibliographical evidence could be found indicating that those drugs may indeed have some potential off-target residual affinity for the nuclear receptors annotated. Overall, the present findings suggest that ligand-based approaches to protein family profiling appear as a promising means toward the establishment of novel tools for in silico pharmacology.

INTRODUCTION

One of the grand challenges in chemical biology is identifying a small-molecule modulator for each individual function of all human proteins.[1] For pharmaceutical research, this has the potential to provide molecules that may then be used as chemical probes for protein validation and as initial hits for lead generation in target and drug discovery programs, respectively.[2] Vital to this aim is the ability to produce quantitative data on the response of biological systems to the presence of chemical compounds.[3] Pharmacologists have been gathering this type of data for over a century. However, it has not been until recently that the technological advances produced in combinatorial chemistry[4] and high-throughput screening[5] have made the collection of these data in a more automatic and systematic manner possible, opening an avenue toward determining experimentally the pharmacological profile of compounds.[6−18] Nonetheless, despite the significant progress made toward improving the capacity for chemical synthesis and particularly for biological testing,[19] any aspiration of being able to make and store every synthetically feasible molecule and test it on every assayable protein remains to date unreachable, and thus, complementary strategies for massive pharmacological profiling of large compound collections need to be explored.[20,21]

One such complementary approach is the application of in silico methods capable of rapidly searching through large virtual chemical spaces for compounds similar to a set of bioactive reference molecules against a panel of multiple targets.[22−24] These methods are based on mathematical representations of molecules[25−29] and capitalize on initiatives aimed at the construction of annotated chemical libraries that incorporate pharmacological data into traditional repositories of chemical structures.[30] Early initiatives focused on the gathering of biological data for drug molecules: of note are the Comprehensive Medicinal Chemistry (CMC) database,[31] offering current biochemical information for over 8400 pharmaceutical compounds, and the Derwent World Drug Index (WDI),[32] containing data on activity and mechanism of action for over 58 000 marketed and development drugs worldwide. More recently, those initiatives have extended their scope to capture the increasing amount of pharmacological data available from public sources. Representative examples are the MDL Drug Data Report (MDDR),[33] including information on therapeutic action and biological activity for over 132 000 compounds gathered from patent literature, journals, and congresses, and the WOMBAT database,[34] offering biological information for 120 400 molecules reported in medicinal chemistry journals over the last 30 years. The construction of all these annotated chemical libraries contribute to the establishment of the knowledge base for integration of chemical and biological data and thus toward a deeper understanding of the properties of molecules

* Corresponding author e-mail: jmestres@imim.es.
† Institut Municipal d'Investigació Mèdica and Universitat Pompeu Fabra.
‡ AstraZeneca.

10.1021/ci600300k CCC: $33.50
Published on Web 09/26/2006

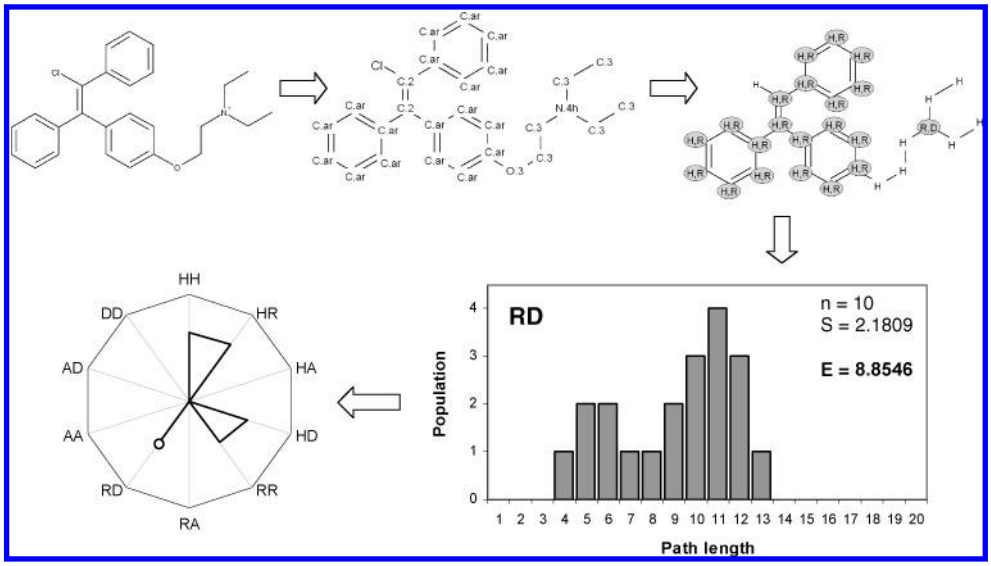

Figure 1. Generation of a SHED profile from chemical structure (see text for details).

associated to the different protein families forming the chemogenomic space.[35,36] Ultimately, the establishment of direct biochemical connections through annotated chemical libraries may contain clues to the existence of apparently unrelated proteins having affinity for similar ligands or the presence of some privileged structures responsible for the activity of ligands in entire protein families.[37,38]

The present work aims at introducing a ligand-based approach to in silico pharmacology by exploiting publicly available pharmacological data collected in a family-directed annotated chemical library encoded using biologically relevant molecular descriptors. In particular, the performance of the approach to the family of nuclear receptors is presented. This family of ligand-activated transcription factors is of paramount importance for pharmaceutical research since many of its members are often considered to be double-edged swords. On one hand, they regulate a variety of biological processes, including lipid and glucose homeostasis, detoxification, cellular differentiation, embryonic development, and organ physiology. Consistent with these important regulatory roles, mutations in nuclear receptors are associated with many common human diseases such as cancer, diabetes, and osteoporosis, and thus, they are considered highly relevant therapeutic targets.[39] On the other hand, there is increasing evidence that nuclear receptors act as regulators of some cytochrome P450 enzymes, which in turn are responsible for the metabolism of molecules. Accordingly, many nuclear receptors are also regarded as potential therapeutic off-targets.[40]

The following sections describe, first, the particular set of molecular descriptors and the annotated chemical library used in this work and, second, the construction of a ligand-based descriptor model as a rapid means for estimating in silico the pharmacological profile of compounds across the family of nuclear receptors.

METHODOLOGY

The use of biologically relevant mathematical representations of molecules and the availability of pharmacological data for a significant number of ligands are the two key elements required to perform the type of analysis presented in this work. Details on the use of Shannon entropy descriptors derived from topological feature−pair distributions and the construction of an annotated chemical library directed to the family of nuclear receptors are provided next.

Shannon Entropy Descriptors. A novel set of molecular descriptors called SHED (Shannon entropy descriptors) was recently introduced.[41] SHED are derived from distributions of atom-centered feature pairs extracted directly from the topology of molecules. The process of obtaining SHED from chemical structure is illustrated in Figure 1 for clomifene, a selective estrogen receptor modulator. The original input structure should be in MDL's SD file format.[42] From a SD file, each atom in a molecule is first mapped to a Sybyl atom type.[43] Subsequently, each atom type is assigned to one or more of four atom-centered features, namely, hydrophobic (H), aromatic (R), acceptor (A), and donor (D). For example, an aliphatic C.3 carbon will be assigned to a hydrophobic feature (H), whereas a protonated N.4h nitrogen will be assigned to both aromatic and donor features (R, D). Then, the shortest path length between atom-centered feature pairs is derived, and its occurrence at different path lengths is stored to create a feature-pair distribution. A maximum path length of 20 bonds was used. Feature pairs at distances over 20 bonds are accumulated in the last bin. As an example, the distribution of RD feature pairs within clomifene is displayed. An equivalent distribution is derived for each of the 10 possible feature pairs resulting from all pair combinations of the four features used.

At this stage, the concept of Shannon entropy is applied to determine the variability of feature-pair distributions.[41] Within this approach, the entropy, S, of a population, P, distributed in a certain number of bins (representing in this case the different path lengths), $N = 20$, is given by

$$S = -\sum_{i=1}^{N} \rho_i \ln \rho_i \qquad \rho_i = p_i/P$$

where ρ_i and p_i are the probability and the population, respectively, at each bin i of the distribution. The values of S range between 0, reflecting the situation of the entire population being concentrated in a single bin, and a

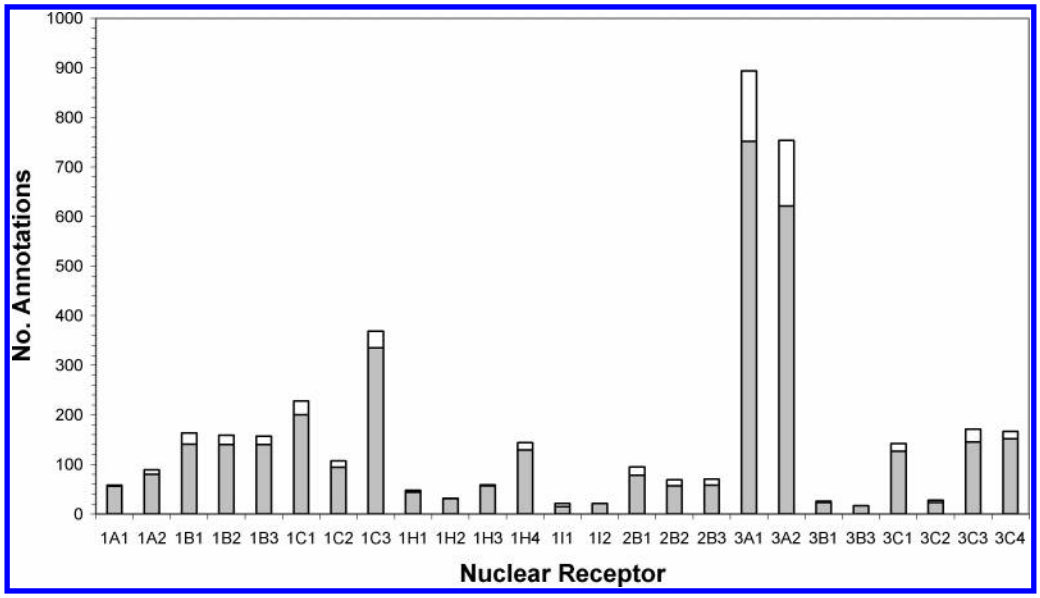

Figure 2. Distribution of all (white bars) and nonredundant (gray bars) chemical annotations present in NRacl among 25 nuclear receptors. See text for details.

maximum number, $S_{max}=\ln N$, reflecting the situation of a uniformly distributed population among all bins. In the case of clomifene (Figure 1), RD pairs can be found at path lengths occupying 10 bins, and the variability in their population results in a distribution with an entropy value of 2.1809. To have a more intuitive measure that can be linearly related to the situation of full uniform occupancy, entropy values are transformed into projected entropy values, $E = e^S$. Correspondingly, the E values provide a measure of the expected maximum uniform occupancy from the corresponding S value. Now, for any given population $P > 0$, the values of E can vary from 1, reflecting the situation of zero entropy in which the population is totally concentrated in a single bin, to N, reflecting the situation of maximum entropy in which the population is uniformly distributed among all bins. In the limit case of $P = 0$, then E will be assigned to $E = 0$. For the RD feature pair in clomifene (Figure 1), the maximum achievable E value for a population uniformly occupying 10 bins would be $E = 10$. The obtained E value of 8.8546 reflects a slight deviation from the situation of full uniform occupancy on 10 bins. This E value will ultimately be the Shannon entropy descriptor (SHED) for the RD feature pair. The set of SHED values obtained for the 10 possible feature pairs constitute the SHED profile of a molecule. As illustrated in Figure 1, SHED profiles can be represented using a wheel chart, the circle in the chart indicating the E value (SHED) for the RD feature pair in clomifene.

Annotated Chemical Library. An annotated chemical library directed to the nuclear receptor family (NRacl) was recently constructed.[44] All data incorporated in NRacl were collected from public sources of information, mainly reviews and medicinal chemistry journals of the last 10 years. Each chemical entity within NRacl contains a set of structural, biological, and bibliographical data. Structural data include a unique identifier for the molecule and its 2D structure representation. Biological data contain the list of nuclear receptors at which the molecule has been reported to be active, identified by their names and corresponding nuclear receptor code, together with the associated pharmacological data (K_i, IC_{50}, and EC_{50}), when available. Finally, the bibliographical data collect the list of references from which structural and biological data were extracted.

A compound in NRacl is considered to be annotated to a given nuclear receptor if its associated pharmacological data (K_i, IC_{50}, and EC_{50}) is under a certain cutoff. In this work, an annotation cutoff of 10 μM was considered. Under this cutoff, NRacl contains currently 4088 annotations to 25 nuclear receptors derived from a total of 2324 molecules, some molecules containing multiple annotations to nuclear receptors. The overall distribution of annotations among all nuclear receptors currently covered by NRacl is provided in Figure 2 (white bars). As can be observed, this distribution is a fair reflection of the historical therapeutic relevance of some of the members of this family. For example, the nuclear receptor containing the largest number of chemical annotations is the estrogen receptor subtype α (ERα; NR3A1), an important target in reproductive medicine and cancer research. Because of its high homology, many compounds binding to ERα are also reported to be active to ERβ, thus justifying the large number of annotations present also for the latter. Another nuclear receptor highly populated with annotations in NRacl is the peroxisome proliferator-activated receptor γ (PPARγ; NR1C3), widely recognized as a key regulator in multiple metabolic pathways including fatty acid and carbohydrate metabolism, and thus, it is considered to be a relevant target in cardiovascular research. In contrast, estrogen-related receptor γ (ERRγ; NR3B3) and vitamin D3 (VDR; NR1I1) are the only nuclear receptors with less than 25 annotations.

All molecules present in NRacl are unique with respect to their structure. However, in terms of their feature distribution, some molecules may have exactly the same SHED profile, either because of pharmacophorically comparable atom mutations or to topologically equivalent structural symmetries (e.g., stereoisomerism, cis/trans isomerism, or symmetric regioisomerism). For example, having a methyl in a molecule substituted by a chlorine in another will give rise to equivalent SHED profiles since the atom types C.3 and Cl are both assigned to a hydrophobic feature.

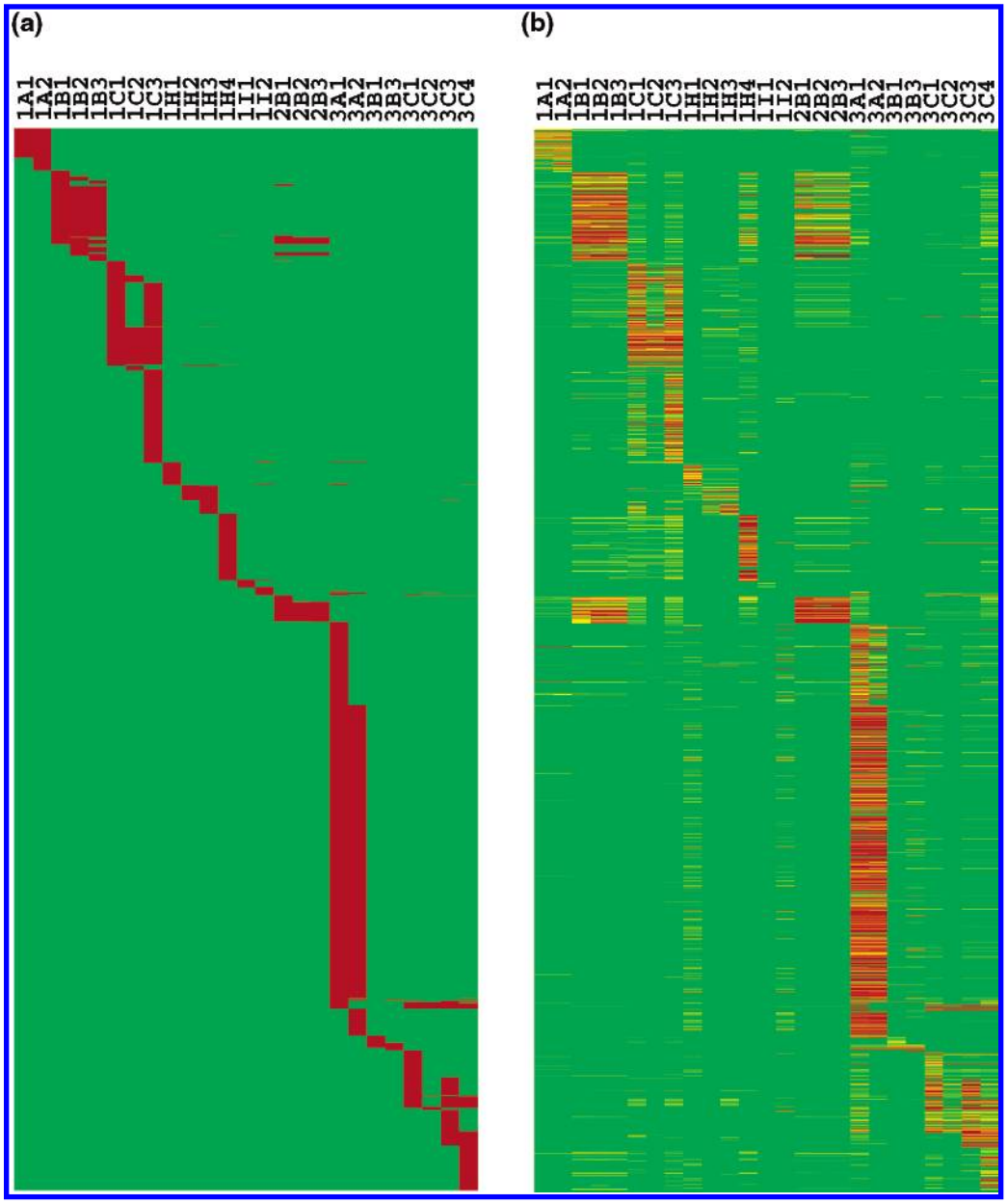

Figure 3. Comparison between the heatmap representing all original annotations extracted from bibliographical sources and stored in NRacl (a) and the heatmap reflecting the minimum SHED Euclidean distances between the SHED profile of each molecule and the set of nonredundant SHED profiles annotated to each nuclear receptor (b). Color coding: (a) red is annotated and green not annotated; and (b) red reflects distance values close to 0.0 and as distances increase in magnitude they turn to orange, yellow, and finally green at a value of 1.2 and over.

Accordingly, to avoid having descriptor collisions that could bias the significance of any statistics derived subsequently, all SHED-redundant molecules were identified and removed. The final distribution among nuclear receptors of the 3536 annotations remaining from the set of 2033 molecules with nonredundant SHED profiles is also depicted in Figure 2 (gray bars).

RESULTS AND DISCUSSION

The distribution of annotations presently contained in NRacl is visually illustrated in Figure 3a. In the heatmap shown, annotations of molecules (in rows) to nuclear receptors (in columns) are represented as red cells, meaning that the interaction of a particular molecule with a specific nuclear receptor has been positively reported and experimentally quantified in the literature with a pharmacological value below 10 μM. In contrast, green cells indicate current lack of information on the possibility of any interaction between a given molecule and a certain nuclear receptor. The extent of the green area denotes the existence of large information gaps, clearly one of the main limitations of dealing with annotated chemical libraries relying on data extracted directly from public sources of information.[44] This is the result of the molecules usually not being screened systematically through a large panel of protein targets, because of limited time and resources, for the sake of obtaining the maximum amount of information possible but instead being screened solely to the target of interest at that point in time. But even if they were screened through multiple targets, habitually only a limited amount of data is made available, since publishing large amounts of negative data is often regarded as not informative. These important, yet often overlooked,

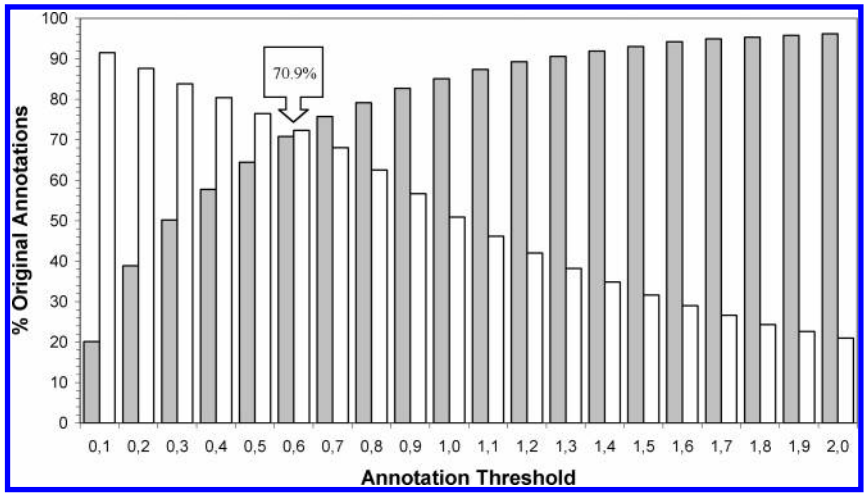

Figure 4. Comparison between the percentage of original annotations in NRacl recovered (gray bars) and the percentage of original annotations within all annotations assigned at each annotation threshold (white bars). A value of 0.6 is taken as a threshold of nuclear receptor concern.

aspects lead to the situation of data incompleteness within the interaction matrix depicted as a heatmap in Figure 3a.

In an attempt to address these limitations, the following section describes a means of filling the gaps in annotated chemical libraries, based on deriving ligand-based descriptor models of protein targets. Subsequently, the applicability of these models for the nuclear receptor profiling of chemical libraries is finally tested on a drugs database and a series of targeted libraries designed for proteases, kinases, ion channels, and G protein-coupled receptors.

Ligand-Based Models of Proteins. The ensemble, S, of nonredundant SHED profiles, s_i^I, representing all molecules, $i = 1,M_I$, annotated to each particular nuclear receptor, $I = 1,N$, constitutes a mathematical description of the nuclear receptor family from a ligand perspective

$$S = [\{s_i^I\}_{i=1,M_I}]_{I=1,N}$$

The scoring of each compound in a chemical library, d^I, with respect to a given nuclear receptor, I, is then assigned to the minimum value of all Euclidean distances calculated between the SHED profile of the compound, s, and each one of the SHED profiles, s_i^I, describing the molecules annotated to that particular nuclear receptor

$$d^I = \text{MINIMUM}\left(\sqrt{\sum_{f=1}^{10}(s_f - \{s_i^I\}_f)^2}\right)_{i=1,M_I}$$

where $\{s_i^I\}_f$ is the SHED corresponding to the f feature−pair distribution of molecule i annotated to nuclear receptor I. In the context of similarity-based virtual screening, the approach of combining the scores over multiple bioactive reference molecules has been recently referred to as group fusion and proven to give significantly superior results to using data fusion strategies on single reference molecules for a wide variety of protein targets.[45]

At this stage, the recall (proportion of original, bibliographically confirmed, annotations stored in NRacl) obtained at different minimum SHED Euclidean distance cutoffs was investigated with the ultimate aim of identifying the optimum cutoff value to be used as the annotation threshold for the nuclear receptor profiling of chemical libraries. Accordingly, the calculation of all minimum Euclidean distances was performed between the SHED profile of each molecule in NRacl and the SHED profiles representing each nuclear receptor. If the latter contained the SHED profile of the molecule in NRacl being processed, that SHED profile was left out during the calculation of Euclidean distances. The results of this analysis are depicted in Figure 4 in which gray bars are the percentage of original annotations recovered and white bars are the percentage of original annotations within all annotations assigned at each minimum SHED Euclidean distance cutoff.

As can be observed in Figure 4, as the annotation threshold is set to higher minimum SHED distances, a larger percentage of original annotations are being recovered (gray bars), but concurrently, they also represent a smaller percentage of all the annotations being assigned to molecules (white bars). An optimal annotation threshold would be one that recovers the majority of the original annotations without adding at the same time too many additional annotations. A minimum SHED Euclidean distance of 0.6 seems to show a satisfactory balance between these two criteria. Under this annotation threshold, a total of 1441 molecules, that is 70.9% of all molecules in NRacl, receive an annotation to at least one nuclear receptor. These 1441 molecules contain a total of 3462 annotations to nuclear receptors, of which 2503 were already present in NRacl and can thus be bibliographically confirmed. These 2503 nuclear receptor annotations represent 70.8% of all annotations present in NRacl (gray bar) and 72.3% of all annotations assigned at this cutoff value (white bar). Since assigning a particular annotation to a molecule effectively reflects a probability for that molecule of having an affinity value under 10 μM for the corresponding nuclear receptor, a minimum SHED Euclidean distance of 0.6 will be considered for the remainder of this work as a threshold of nuclear receptor concern when profiling chemical libraries. This strategy follows on recent studies suggesting that similarity to molecules in the reference set is a good criteria for prediction accuracy of external test sets.[46]

As an illustrative example, Figure 5 shows the profile of minimum SHED Euclidean distances obtained for molecule **1** over the ligand-based model of 25 nuclear receptors. On the basis of literature data, molecule **1** is annotated in NRacl to the three retinoic acid receptors (RARα, 1B1; RARβ, 1B2;

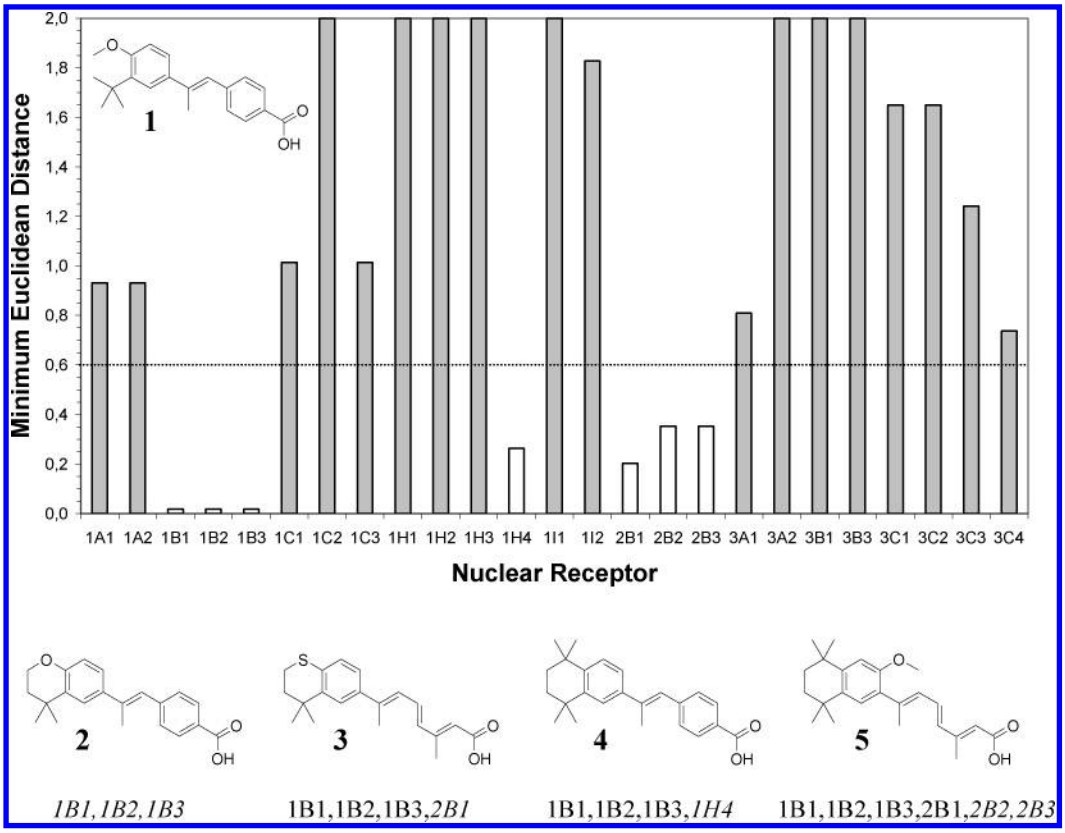

Figure 5. Nuclear receptor profile of molecule **1** based on the minimum SHED Euclidean distance between its SHED profile and the set of nonredundant SHED profiles annotated to each nuclear receptor. Molecules **2**−**5** are the molecules present in NRacl responsible for the annotations (white bars) assigned to molecule **1**.

and RARγ, 1B3).[47] However, as reflected in Figure 5, under the 0.6 annotation threshold established above, molecule **1** can be annotated to seven nuclear receptors (white bars). The lowest values of all minimum Euclidean distances correspond indeed to the three RAR nuclear receptors, thus recovering all annotations assigned originally to molecule **1**. In addition to the RAR annotations, Euclidean distances between 0.20 and 0.35 annotate molecule **1** to the three retinoic X receptors (RXRα, 2B1; RXRβ, 2B2; and RXRγ, 2B3) and the farnesoid X receptor (FXR, 1H4).

One of the advantages of using a ligand-based approach for annotating molecules to nuclear receptors is the possibility of examining the ligands responsible for the annotations and, if necessary, of going back to the original bibliographical sources stored in NRacl. The four molecules responsible for the seven annotations to molecule **1** are also collected in Figure 5. The full list of annotations assigned to each of those molecules in NRacl is also included and those annotations furnished by each respective molecule to molecule **1** are given in italics. Among all molecules currently present in NRacl, molecule **2** is, with a SHED Euclidean distance of 0.0177, the most similar molecule to molecule **1**. It has been reported to have potencies below 10 μM for 1B1, 1B2, and 1B3[48] and is thus responsible for the assignment of those annotations to molecule **1**. Also, molecule **3** has been reported to have potencies below 10 μM for 1B1, 1B2, 1B3, and 2B1[48] and, with a SHED Euclidean distance of 0.2033, is responsible for the 2B1 annotation to molecule **1**. In addition, molecule **4** (also known as TTNPB) was the first nonsteroidal ligand to be described and is known to be a weak FXR agonist.[49] At a SHED Euclidean distance of 0.2630, molecule **4** is ultimately responsible for the 1H4 annotation to molecule **1**. Finally, molecule **5** has been reported to have binding affinities below 10 μM for all RARs and RXRs.[50] With a SHED Euclidean distance of 0.3527, molecule **5** is responsible for the additional 2B2 and 2B3 annotations to molecule **1**. In addition to the reported potencies for the three RARs, we are not able to confirm that molecule **1** indeed has affinities below 10 μM for FXR and the three RXRs, but the evident structural similarities with the four molecules responsible for all annotations are an indication that those additional four annotations are not an unreasonable alert.

The process described above for molecule **1** was applied to each one of the 2033 molecules with nonredundant SHED profiles present in NRacl. The results are given in Figure 3b, in which the order of the molecules is exactly the same as the one obtained from the original annotations shown in Figure 3a. In contrast to the binary heatmap illustrated in Figure 3a, in which red was annotated and green was not annotated, Figure 3b presents a color gradation between red and green reflecting the value of the minimum SHED Euclidean distance between the SHED profile of each molecule and the set of nonredundant SHED profiles annotated to each nuclear receptor. With the annotation threshold of 0.6 taken as the center of the color scale, distance values close to 0.0 are represented in red, those close to 0.6 are seen as light orange, and as the distances increase in magnitude they turn to yellow and finally green at a value of 1.2 and over. There are two main aspects worth mentioning when comparing the heatmaps of Figure 3a and b. On one hand, it is remarkable the fact that the essential pattern observed when plotting the original annotations (Figure 3a) is preserved when molecules are processed through the

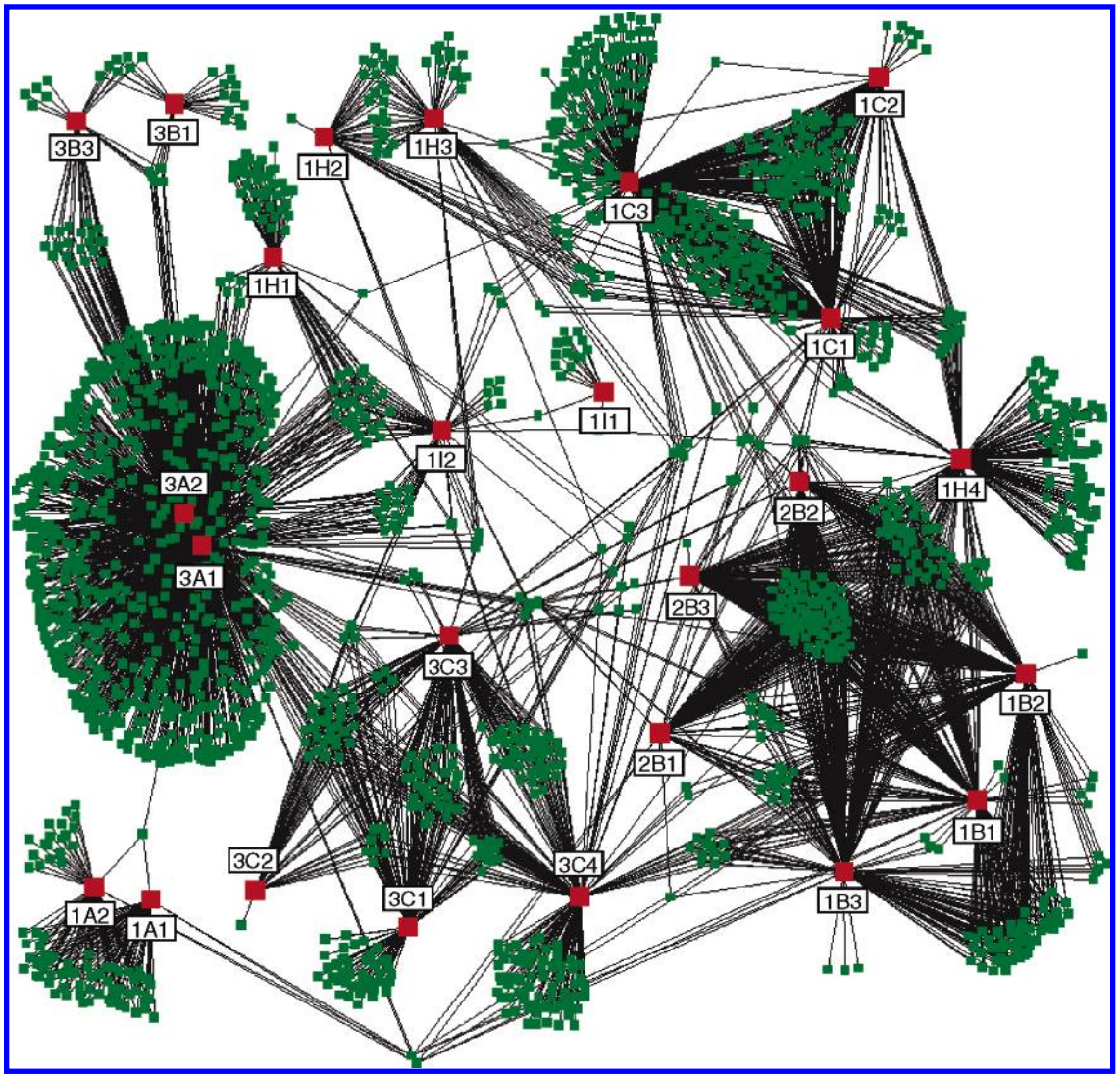

Figure 6. Nuclear receptor network derived using the matrix of minimum SHED Euclidean distances (see Figure 3b). The set of 2033 molecules from NRacl are given as small green squares and the 25 nuclear receptors as large red squares. A value of 0.6 for the minimum SHED Euclidean distance was used to establish direct links between molecules and nuclear receptors. This linkage map was constructed with Cytoscape.[51]

ligand-based descriptor model of nuclear receptors (Figure 3b). This result reveals that the remaining molecules in NRacl are to a great extent representative of the molecule being processed after leaving that molecule out; something that can only be achieved if the annotated chemical space has been sufficiently saturated with as many known bioactive molecules as possible. On the other hand, despite the clear discrimination between nuclear receptor groups, some correlation patterns between them emerge. The most apparent example is the clear correlation observed in Figure 3b between RARs (1B1, 1B2, and 1B3) and RXRs (2B1, 2B2, and 2B3): a result that provides an indication of the potential of this approach for understanding side effects through the identification of off-target affinities.

To further investigate any potential links between nuclear receptors from the perspective of their active ligands, we used some graph-based tools to construct an interaction network.[51] A similar method was reported recently to visualize nuclear hormone receptor networks relevant to drug metabolism.[52] Figure 6 contains the nuclear receptor network obtained on the basis of the matrix of minimum SHED Euclidean distances presented in Figure 3b, using the threshold of nuclear receptor concern derived above as the linkage criteria for assigning direct connections between a molecule and a nuclear receptor. The first observation that can be made from inspection of Figure 6 is that the essential phylogenetic relationships among nuclear receptors are essentially preserved. The strong interconnections between RARs and RXRs noticed above are also clearly retrieved. But perhaps a more interesting outcome is the existence of several molecules connecting the estrogen receptors (3A1 and 3A2) with the pregnane X receptor (PXR, 1I2) and the ecdysone receptor (1H1). The link observed between PXR and estrogen may have implications related to the metabolism and toxicity of estrogenic compounds,[53,54] whereas the link identified between ecdysone and estrogen may suggest that a library designed around estrogenic compounds may be a good starting point for the identification of novel chemical modulators of the ecdysone receptor.[55] Finally, notable are the links observed between FXR (1H4) and the peroxisome proliferator-activated receptor γ (PPARγ, 1C3), as well as with members of the RAR and RXR groups, indicating that FXR may be an off-target to take into consideration when developing compounds for PPARγ, RARs, and RXRs.[56,57]

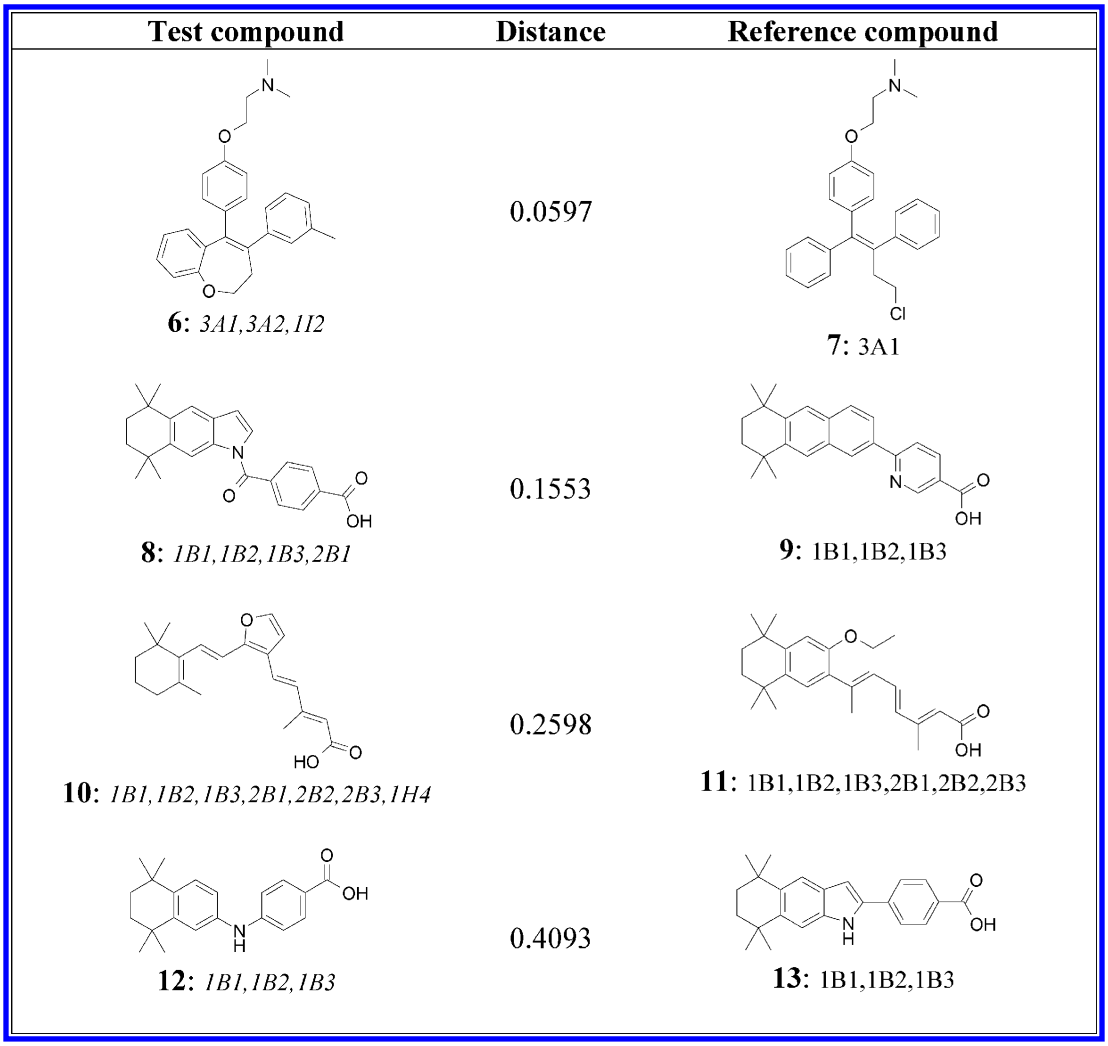

Figure 7. Selection of molecules from the test set annotated to nuclear receptors (test compound). The molecule in NRacl showing the minimum SHED Euclidean distance to each test compound (reference compound) is also included. In addition to the annotations provided specifically by the reference compound shown, all annotations assigned to test compounds from reference compounds are also given in italics (see Figure 5).

Nuclear Receptor Profiling. The ligand-based model of nuclear receptors described in the previous section can then be used to profile large chemical libraries of this family of transcription factors of key importance for pharmaceutical research. As a first validation exercise, a small database of 82 diverse molecules, not included in NRacl, was compiled as an external test set from bibliographic sources reporting experimental evidence of activity on nuclear receptors. The profiling of this database on the ligand-based nuclear receptor model provided annotations (that is, a minimum SHED Euclidean distance of 0.6 for at least one nuclear receptor) for 47 molecules (57% of the database). Out of this 47 molecules, 32 molecules had annotations to at least one member of the nuclear receptor group at which the molecules were known to be active. Accordingly, the method provided a correct identification of the target nuclear receptor group for 68.1% of the molecules annotated. This result is in good agreement with the expected performance of the method (70.9%, vide supra).

To provide an idea of the level of structural hopping that can be achieved using the current approach, a selection from this 47 molecules, covering the whole range of distance values under the threshold of 0.6, is presented in Figure 7. Molecule **6** is a benzoxepin analogue of tamoxifen.[58] Not surprisingly, the closest molecule found when profiled against the ligand-based nuclear receptor model is molecule **7**, toremifene, a chlorine analogue of tamoxifen, annotated in NRacl to have affinity for the estrogen receptor subtype α (ERα, 3A1), and thus it is responsible for the 3A1 annotation to molecule **6**.[59] Molecule **6** also had minimum SHED Euclidean distances below 0.6 for other molecules in NRacl that provided additional annotations to the estrogen receptor subtype β (ERβ, 3A2) and PXR (1I2). Molecule **8**, Am93, is structurally related to Am80, a known potent synthetic retinoid reported to be active to RARs.[60] Nuclear receptor profiling of this compound identifies molecule **9**, with annotations to all RARs,[61] as the molecule in NRacl being closest in SHED profile to molecule **8**. Molecule **10** is a conformationally restrained analogue of retinoic acid, with reported nanomolar affinity for all RXRs and 100-fold selectivity over RARs.[62] Molecule **11** is the closest reference compound found to it in NRacl. Most interestingly, despite being annotated to all RARs and RXRs because of the 10 μM cutoff used for the construction of NRacl, molecule **11** is also being reported to have nanomolar affinities for all RXRs but only micromolar affinities for all RARs, in good agreement with the known profile of molecule **10**.[50] Finally, molecule **12**, DA010, is a known ligand for both RARs and RXRs.[63] In this case, molecule **13** is the only molecule present in NRacl with a

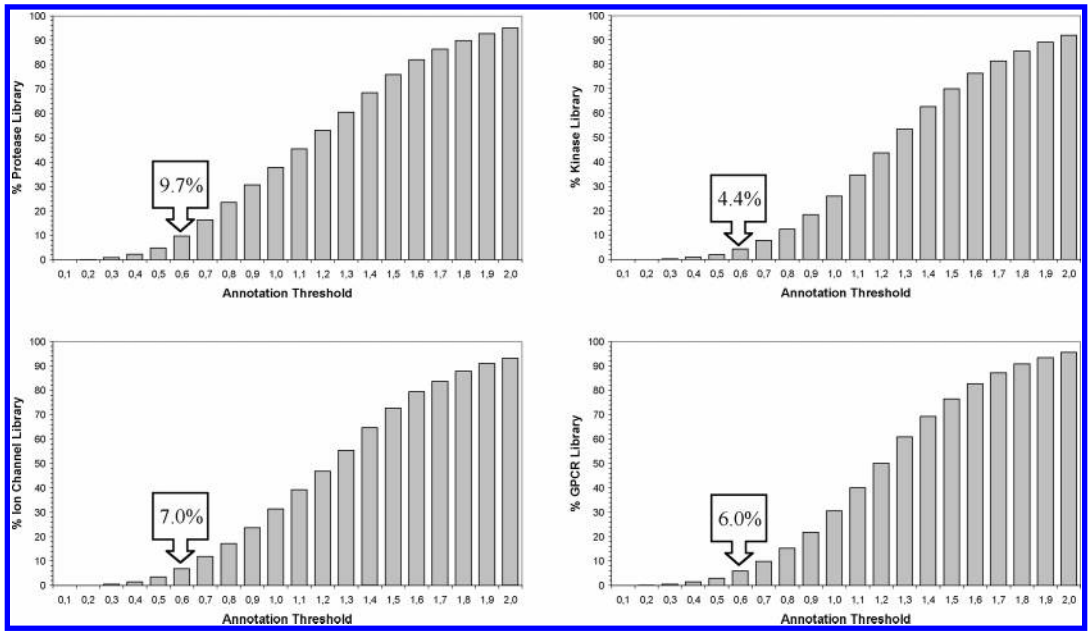

Figure 8. Percentage of each targeted library (protease, kinase, ion channel, and GPCR) annotated to nuclear receptors at varying threshold values. Numbers indicate the percentage of each library under the threshold of nuclear receptor concern.

SHED profile under a distance of 0.6 which is similar to molecule **12** and is thus responsible for the annotation of molecule **12** to all RARs.[61] This is an example where the approach would have missed annotating molecule **12** to all RXRs.

A direct application of this approach, following current trends in chemogenomic strategies for family directed drug discovery, is in the selection of compounds targeted to nuclear receptors from chemical provider databases and the nuclear receptor profiling of targeted libraries designed specifically for other protein families of high therapeutic relevance. Accordingly, as a second validation exercise, we took four commercially available targeted libraries designed to provide hits for the families of proteases, kinases, ion channels, and G protein-coupled receptors (GPCR) containing 19 649, 515 265, 31 579, and 110 507 molecules, respectively.[64] For the sake of consistency, a set of 10 000 molecules was randomly selected from each of them and processed through the ligand-based nuclear receptor model. The results are compiled in Figure 8, represented in a form comparable to the gray bars in Figure 4. As can be observed, under the annotation threshold of 0.6 established above, only 9.7% of the 10 000 molecules contained in the protease library, 4.4% of an equivalent number of molecules present in the kinase library, 7.0% of the ion channel library, and 6.0% of the GPCR library would be alerted on the potential of having affinity to some nuclear receptor. Since molecules designed for these libraries should not be expected to have activity on nuclear receptors, at least to a large extent, these results confirm the validity of the ligand-based approach to the nuclear receptor profiling used in this work.

One of the major concerns during the process of drug discovery is the possibility that the compounds being optimized could have residual affinities for some off-targets responsible for highly undesirable side effects. Therefore, the third validation exercise consisted of profiling a dataset of 2944 drugs, none of them present in NRacl, through the ligand-based nuclear receptor model. Interestingly, only 153 drugs, that is 5.2% of the total number contained in the dataset, were identified as having at least a minimum SHED Euclidean distance below 0.6 to a nuclear receptor. Of those, 32 drugs contained a steroidal scaffold and were found to be similar to some of the steroidal hormone receptor ligands present in NRacl. An additional set of 19 drugs was found to have the same scaffold to some molecule present in NRacl and could thus be considered to be close analogues. An illustrative selection of the 102 remaining drugs that were annotated to nuclear receptors is compiled in Figure 9, covering the whole range of distances within the 0.6 annotation threshold.

The first drug collected in Figure 9 is acitretin, an oral retinoid indicated for the treatment of psoriasis, although its mechanism of action has not been fully elucidated. When profiled against the ligand-based nuclear receptor model, this drug was found at close distance to molecule **14**, a molecule that has been reported to have potencies below 10 μM to all three RARs.[47] This result would provide an alert for acitretin having potential affinity for RARs. We have not been able to find direct experimental data confirming this fact, but there is bibliographical evidence suggesting a link between psoriasis and an alteration in the cellular retinoid pathways at which synthetic retinoids, such as acitretin, may interact.[65] The next drug listed is telmisartan, originally designed as an angiotensin II receptor blocker for treating the metabolic syndrome. Nuclear receptor profiling of this drug identified molecule **15** as the closest to its SHED profile, a molecule that has been reported to have affinity below 10 μM to all peroxisome proliferator-activated receptor subtypes (PPARα, 1C1; PPARβ, 1C2; PPARγ, 1C3).[66] Recent preclinical studies indicate that telmisartan acts as a PPARγ modulator when tested at concentrations that might be achievable with oral doses recommended for the treatment of hypertension.[67] If telmisartan had been processed through some type of nuclear receptor alert system, such as the one presented in this work, this outcome could have been anticipated at an earlier stage. The next drug is phenyltoloxamine, a H1 histamine antagonist used in pain treatment. Molecule **16** was identified as the closest molecule in NRacl to this drug. This molecule has been reported to show antagonistic activity

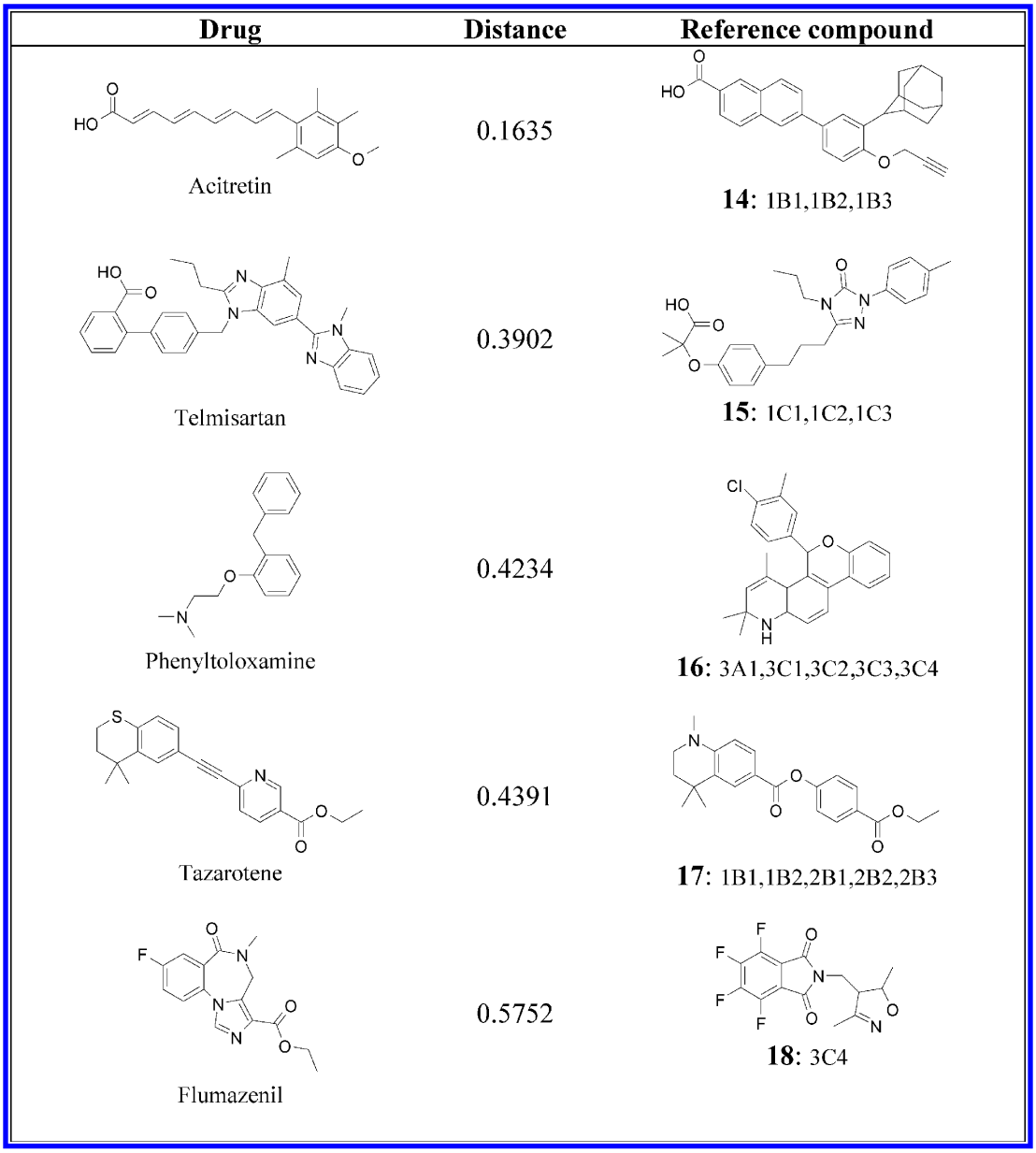

Figure 9. Selection of drugs annotated to nuclear receptors yet not designed specifically for them. The molecule in NRacl showing the minimum SHED Euclidean distance to each drug (reference compound) is also included.

for ERα (3A1), as well as to the glucocorticoid (GR, 3C1), mineralcorticoid (MR, 3C2), progesterone (PR, 3C3), and androgen (AR, 3C4) receptors.[68] Because the SHED Euclidean distance between phenyltoloxamine and molecule **16** is below 0.6, phenyltoloxamine is alerted for having potential affinity on these five nuclear receptors. We have not been able to find experimental data confirming this result, but bibliographical evidence has been found on the fact that the antiestrogen binding site is a histamine or histamine-like receptor and that, in particular, phenyltoloxamine has indeed some antagonistic affinity for the estrogen receptor.[69] Next in Figure 7 is tazarotene, a topically applied retinoid for the treatment of psoriasis. Nuclear receptor profiling of this drug identifies molecule **17**, with annotations to RARs and RXRs, as the molecule in NRacl being closest in SHED profile to tazarotene.[70] Indeed, indications can be found in the literature that tazarotene does bind to RARs, although selectively with respect to RXRs.[71] Finally, flumazenil is a $GABA_A$-benzodiazepine receptor antagonist. A SHED Euclidean distance right below 0.6 is obtained with respect to molecule **18**, a known androgen receptor antagonist.[72] Despite the fact that no direct evidence of this result could be found, recent indirect reports indicate that flutamide, an androgen receptor antagonist, produced an anticonvulsant effect in common seizure models through a possible interaction with benzodiazepine receptors, thus being indicative of potential cross-pharmacologies between these two receptors.[73]

CONCLUSIONS

On the basis of pharmacological data extracted directly from bibliographical sources, a ligand-based descriptor model for the family of nuclear receptors was constructed, offering the possibility to perform in silico profiling of large chemical libraries on 25 nuclear receptors in a fast and efficient manner. It was shown that, provided that the annotated chemical space for the protein family of interest is sufficiently well saturated, the model attains a decent degree of both internal consistency and external predictability. The model served also to construct an interaction network from which potential cross-pharmacologies between nuclear receptors

emerged. In addition, the approach was proven to be sensible enough to achieve significant discriminative power when applied to external chemical libraries designed for a priori unrelated protein families such as proteases, kinases, ion channels, and GPCRs, opening an avenue for its use in the selection and design of targeted libraries. Further external validation of the model was finally provided by the identification of a selected list of drugs for which bibliographical evidence exists, though indirect for some, indicating that those drugs may indeed have some potential off-target residual affinity for the nuclear receptors annotated.

A vast amount of pharmacological data on many protein targets is becoming available for increasingly large quantities of molecules. The systematic collection of these data in annotated chemical libraries allows for describing proteins from the perspective of ligands, descriptions that develop into a more complete picture as data on new ligands are obtained and the active chemical space stored becomes more saturated. Interestingly, a recent study comparing the performance of some ligand-based and structure-based methods for virtual screening concluded that information about a target derived from knowledge on bioactive ligands can be as valuable as knowledge of the target structures for identifying novel scaffolds by computational means.[74] Functional coverage of the proteome by structures is progressing rapidly but many areas are still devoid of any structural information.[75] In the wait for at least one representative structure for each target protein, ligand-based representations of proteins may offer a means to move from the traditional virtual chemical screening to the necessary virtual pharmacological profiling.

The ligand-based approach to nuclear receptor profiling presented in this work can be readily extended to other protein families of high-therapeutic relevance, such as GPCRs, for which only limited structural information is available but pharmacological data on a significant number of molecules are known. Via the gathering, proper storing, and maximal exploitation of all pharmacological knowledge on ligands available to date, in silico pharmacology on a genomic scale may soon become a reality.

ACKNOWLEDGMENT

This research was funded by the European Commission (InfoBioMed network of excellence IST-507585), the Spanish Ministerio de Educación y Ciencia (project reference BIO2005-04171), and the Instituto de Salud Carlos III.

REFERENCES AND NOTES

(1) Schreiber, S. L. Small Molecules: The Missing Link in the Central Dogma. *Nature Chem. Biol.* **2005**, *1*, 64−66.
(2) Bredel, M.; Jacoby, E. Chemogenomics: An Emerging Strategy for Rapid Target and Drug Discovery. *Nature Rev. Genet.* **2004**, *5*, 262−275.
(3) Walters, W. P.; Namchuk, M. Designing Screens: How to Make your Hits a Hit. *Nature Rev. Drug Discovery* **2003**, *2*, 259−266.
(4) Ramström, O.; Lehn, J.-M. Drug Discovery by Dynamic Combinatorial Libraries. *Nature Rev. Drug Discovery* **2002**, *1*, 26−36.
(5) Sundberg, S. A. High-Throughput and Ultrahigh-Throughput Screening: Solution- and Cell-Based Approaches. *Curr. Opin. Biotechnol.* **2000**, *11*, 47−53.
(6) Kauvar, L. M.; Higgins, D. L.; Villar, H. O.; Sportsman, J. R.; Engqvist-Goldstein, A.; Bukar, R.; Bauer, K. E.; Dilley, H.; Rocke, D. M. Predicting Ligand Binding to Proteins by Affinity Fingerprinting. *Chem. Biol.* **1995**, *2*, 107−118.
(7) Weinstein, J. N.; Myers, T. G.; O'Connor, T. M.; Friend, S. H.; Fornace, A. J., Jr.; Kohn, K. W.; Fojo, T.; Bates, S. E.; Rubinstein, L. V.; Anderson, N. L.; Buolamwini, J. K.; van Osdol, W. W.; Monks, A. P.; Scudiero, D. A.; Sausville, E. A.; Zaharevitz, D. W.; Bunow, B.; Viswanadhan, V. N.; Johnson, G. S.; Wittes, R. E.; Paull, K. D. An Information-Intensive Approach to the Molecular Pharmacology of Cancer. *Science* **1997**, *275*, 343−349.
(8) MacBeath, G.; Koehler, A. N.; Schreiber, S. L. Printing Small Molecules as Microarrays and Detecting Protein−Ligand Interactions en Masse. *J. Am. Chem. Soc.* **1999**, *121*, 7967−7968.
(9) Poulain, R.; Horvath D.; Bonnet, B.; Eckhoff, C.; Chapelain, B.; Bodinier, M.-C.; Déprez, B. From Hit to Lead: Analyzing Structure−Profile Relationships. *J. Med. Chem.* **2001**, *44*, 3391−3401.
(10) Rabow, A. A.; Shoemaker, R. H.; Sausville, E. A.; Covell, D. G. Mining the National Cancer Institute's Tumor Screening Database: Identification of Compounds with Similar Cellular Activities. *J. Med. Chem.* **2002**, *45*, 818−840.
(11) Greenbaum, D. C.; Arnold, W. D.; Lu, F.; Hayrapetian, L.; Baruch, A.; Krumrine, J.; Toba, S.; Chehade, K.; Brömme, D.; Kuntz, I. D.; Bogyo, M. Small Molecule Affinity Fingerprinting: A Tool for Enzyme Family Subclassification, Target Identification, and Inhibitor Design. *Chem. Biol.* **2002**, *9*, 1085−1094.
(12) Kunkel, E. J.; Dea, M.; Ebens, A.; Hytopoulos, E.; Melrose, J.; Nguyen, D.; Ota, K. S.; Plavec, I.; Wang, Y.; Watson, S. R.; Butcher, E. C.; Berg, E. L. An Integrative Biology Approach for Analysis of Drug Action in Models of Human Vascular Inflammation. *FASEB J.* **2004**, *18*, 1279−1281.
(13) Giaever, G.; Flaherty, P.; Kumm, J.; Proctor, M.; Nislow, C.; Jaramillo, D. F.; Chu, A. M.; Jordan, M. I.; Arkin, A. P.; Davis, R. W. Chemogenomic Profiling: Identifying the Functional Interactions of Small Molecules in Yeast. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 793−798.
(14) Szakács, G.; Annereau, J.-P.; Lababidi, S.; Shankavaram, U.; Arciello, A.; Bussey, K. J.; Reinhold, W.; Guo, Y.; Kruh, G. D.; Reimers, M.; Weinstein, J. N.; Gottesman, M. M. Predicting Drug Sensitivity and Resistance: Profiling ABC Transporter Genes in Cancer Cells. *Cancer Cell* **2004**, *6*, 129−137.
(15) Roth, B. L.; Sheffler, D. J.; Kroeze, W. K. Magic Shotguns versus Magic Bullets: Selectively Non-Selective Drugs for Mood Disorders and Schizophrenia. *Nature Rev. Drug Discovery* **2004**, *3*, 353−359.
(16) Fabian, M. A.; Biggs, W. H., III; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lélias, J.-M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. A Small Molecule−Kinase Interaction Map for Clinical Kinase Inhibitors. *Nature Biotechnol.* **2005**, *23*, 329−336.
(17) Fliri, A. F.; Loging, W. T.; Thadeio, P. F.; Volkmann, R. A. Analysis of Drug Induced Effect Patterns to Link Structure and Side Effects of Medicines. *Nature Chem. Biol.* **2005**, *1*, 389−397.
(18) Ramanathan, A.; Wang, C.; Schreiber, S. L. Perturbational Profiling of a Cell-Line Model of Tumorigenesis by Using Metabolic Measurements. *Proc. Natl. Acad. Sci.* **2005**, *102*, 5992−5997.
(19) Haiching M.; Horiuchi, K. Y. Chemical Microarray: A New Tool for Drug Screening and Discovery. *Drug Discovery Today* **2006**, *11*, 661−668.
(20) Bajorath, J. Integration of Virtual and High-Throughput Screening. *Nature Rev. Drug Discovery* **2002**, *1*, 882−894.
(21) Hopkins, A. L.; Mason, J. S.; Overington, J. P. Can We Rationally Design Promiscuous Drugs? *Curr. Opin. Struct. Biol.* **2006**, *16*, 127−136.
(22) Krejsa, C. M.; Horvath, D.; Rogalski, S. L. Penzotti, J. E.; Mao, B.; Barbosa, F.; Migeon, J. C. Predicting ADME Properties and Side Effects: The BioPrint Approach. *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 470−480.
(23) Root, D. E.; Flaherty, S. P.; Kelley, B. P.; Stockwell, B. R. Biological Mechanism Profiling Using an Annotated Compound Library. *Chem. Biol.* **2003**, *10*, 881−892.
(24) Strausberg, R. L.; Schreiber, S. L. From Knowing to Controlling: A Path from Genomics to Drugs Using Small Molecule Probes. *Science* **2003**, *300*, 294−295.
(25) Schuffenhauer, A.; Floersheim, P.; Acklin, P.; Jacoby, E. Similarity Metrics for Ligands Reflecting the Similarity of the Target Proteins. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 391−405.
(26) Horvath, D.; Jeandenans, C. Neighborhood Behavior of In Silico Structural Spaces with Respect to In Vitro Activity Spaces: A Novel Understanding of the Molecular Similarity Principle in the Context of Multiple Receptor Binding Profiles. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 680−690.
(27) Poroikov, V. V.; Filimonov, D. A.; Borodina, Y. V.; Lagunin, A. A.; Kos, A. Robustness of Biological Activity Spectra Predicting by Computer Program PASS for Noncongeneric Sets of Chemical Compounds. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1349−1355.

(28) Cleves, A. E.; Jain, A. N. Robust Ligand-Based Modeling of the Biological Targets of Known Drugs. *J. Med. Chem.* **2006**, *49*, 2921−2938.
(29) Nidhi; Glick, M.; Davies, J. W.; Jenkins, J. L. Prediction of Biological Targets for Compounds Using Multiple-Category Bayesian Models Trained on Chemogenomics Databases. *J. Chem. Inf. Model.* **2006**, *46*, 1124−1133.
(30) Savchuk, N. P.; Balakin, K. V.; Tkachenko, S. E. Exploring the Chemogenomic Knowledge Space with Annotated Chemical Libraries. *Curr. Opin. Chem. Biol.* **2004**, *8*, 412−417.
(31) *MDL Comprehensive Medicinal Chemistry (CMC-3D) Database*; MDL Information Systems, Inc: San Leandro, CA. CMC is an updated electronic version of *Comprehensive Medicinal Chemistry*; Pergamon Press: Elmsford, NY, 1990; Vol. 6.
(32) *World Drug Index*; Derwent Information Ltd.: London.
(33) *MDL Drug Data Report*; MDL Information Systems, Inc.: San Leandro, CA.
(34) Olah, M.; Mracec, M.; Ostopovici, L.; Rad, R.; Bora, A.; Hadaruga, N.; Olah, I.; Banda, M.; Simon, Z.; Oprea, T. I. WOMBAT: World of Molecular Bioactivity. In *Chemoinformatics in Drug Discovery*; Oprea, T. I., Ed.; Wiley-VCH: New York, 2004; pp 223−239; Sunset Molecular Discovery, LLC.: Santa Fe, NM, 2004.
(35) Vieth, M.; Sutherland, J. J.; Robertson, D. H.; Campbell, R. M. Kinomics: Characterizing the Therapeutically Validated Kinase Space. *Drug Discovery Today* **2005**, *10*, 839−846.
(36) Vieth, M.; Sutherland, J. J. Dependence of Molecular Properties on Proteomic Family for Marketed Oral Drugs. *J. Med. Chem.* **2006**, *49*, 3451−3453.
(37) Koch, M. A.; Waldmann, H. Protein Structure Similarity Clustering and Natural Product Structure as Guiding Principles in Drug Discovery. *Drug Discovery Today* **2005**, *10*, 471−483.
(38) Klabunde, T.; Hessler, G. Drug Design Strategies for Targeting G Protein-Coupled Receptor. *ChemBioChem* **2002**, *3*, 928−944.
(39) Gronemeyer, H.; Gustafsson, J. A.; Laudet, V. Principles for Modulation of the Nuclear Receptor Superfamily. *Nature Rev. Drug Discovery* **2004**, *3*, 950−964.
(40) Francis, G. A.; Fayard, E.; Picard, F.; Auwerx, J. Nuclear Receptors and the Control of Metabolism. *Annu. Rev. Physiol.* **2003**, *65*, 261−311.
(41) Gregori-Puigjané, E.; Mestres, J. SHED: Shannon Entropy Descriptors from Topological Feature Distributions. *J. Chem. Inf. Model.* **2006**, *46*, 1615−1622.
(42) Elsevier MDL, San Leandro, CA, http://www.mdli.com/ (accessed May 2006).
(43) Tripos, Inc., St. Louis, MO, http://www.tripos.com/mol2/atom_types.html (accessed May 2006).
(44) Cases, M.; García-Serna, R.; Hettne, K.; Weeber, M.; van der Lei, J.; Boyer, S.; Mestres, J. Chemical and Biological Profiling of an Annotated Compound Library Directed to the Nuclear Receptor Family. *Curr. Top. Med. Chem.* **2005**, *5*, 763−772.
(45) Hert, J.; Willett, P.; Wilton, D. J.; Acklin, P.; Azzaoui, K.; Jacoby, E.; Schuffenhauer, A. Comparison of Fingerprint-Based Methods for Virtual Screening Using Multiple Bioactive Reference Structures. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1177−1185.
(46) Sheridan, R. P.; Feuston, B. P.; Maiorov, V. N.; Kearsley, S. K. Similarity to Molecules in the Training Set is a Good Discriminator for Prediction Accuracy in QSAR. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1912−1928.
(47) Douguet, D.; Thoreau, E.; Grassy, G. Quantitative Structure−Activity Relationship Studies of RAR α, *β*, and *γ* Retinoid Agonists. *Quant. Struct.−Act. Relat.* **1999**, *18*, 107−123.
(48) Benbrook, D. M.; Subramanian, S.; Gale, J. B.; Liu, S.; Brown, C. W.; Boehm, M. F.; Berlin, K. D. Synthesis and Characterization of Heteroarotinoids Demonstrate Structure Specificity Relationships. *J. Med. Chem.* **1998**, *41*, 3753−3757.
(49) Maloney, P. R.; Parks, D. J.; Haffner, C. D.; Fivush, A. M.; Chandra, G.; Plunket, K. D.; Creech, K. L.; Moore, L. B.; Wilson, J. G.; Lewis, M. C.; Jones, S. A.; Willson, T. M. Identification of a Chemical Tool for the Orphan Nuclear Receptor FXR. *J. Med. Chem.* **2000**, *43*, 2971−2974.
(50) Canan Koch, S. S.; Dardashti, L. J.; Hebert, J. J.; White, S. K.; Croston, G. E.; Flatten, K. S.; Heyman, R. A.; Nadzan, A. M. Identification of the First Retinoid X Receptor Homodimer Antagonist. *J. Med. Chem.* **1996**, *39*, 3229−3234.
(51) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498−2504.
(52) Ekins, S.; Kirillov, E.; Rakhmatulin, E. A.; Nikolskaya, T. A Novel Method for Visualizing Nuclear Hormone Receptor Networks Relevant to Drug Metabolism. *Drug Metab. Dispos.* **2005**, *33*, 474−481.
(53) Kretschmer, X. C.; Baldwin, W. S. CAR and PXR: Xenosensors of Endocrine Disrupters? *Chem. Biol. Interact.* **2005**, *155*, 111−128.
(54) Ding, X.; Lichti, K.; Staudinger, J. L. The Micoestrogen Zearalenone Induces CYP3A Through Activation of the Pregnane X Receptor. *Toxicol. Sci.* **2006**, *91*, 448−455.
(55) Oberdorster, E.; Clay, M. A.; Cottam, D. M.; Wilmot, F. A.; McLachlan, J. A.; Milner, M. J. Common Phytochemicals are Ecdysteroid Agonists and Antagonists: A Possible Evolutionary Link Between Vertebrate and Invertebrate Steroid Hormones. *J. Steroid. Biochem. Mol. Biol.* **2001**, *77*, 229−238.
(56) Fiorucci, S.; Rizzo, G.; Antonelli, E.; Renga, B.; Mencarelli, A.; Riccardi, L.; Morelli, A.; Pruzanski, M.; Pellicciari, R. Cross-talk between Farnesoid-X-Receptor (FXR) and Peroxisome Proliferator-Activated Receptor *γ* Contributes to the Antifibrotic Activity of FXR Ligands in Rodent Models of Liver Cirrhosis. *J. Pharmacol. Exp. Ther.* **2005**, *315*, 58−68.
(57) Kassam, A.; Miao, B.; Young, P. R.; Mukherjee, R. Retinoid X Receptor (RXR) Agonist-Induced Antagonism of Farnesoid X Receptor (FXR) Activity Due to Absence of Coactivator Recruitment and Decreased DNA Binding. *J. Biol. Chem.* **2003**, *278*, 10028−10032.
(58) Lloyd, D. G.; Hughes, R. B.; Zisterer, D. M.; Williams, D. C.; Fattorusso, C.; Catalanotti, B.; Campiani, G.; Meegan, M. J. Benzoxepin-Derived Estrogen Receptor Modulators: A Novel Molecular Scaffold for the Estrogen Receptor. *J. Med. Chem.* **2004**, *47*, 5612−5615.
(59) Fang, H.; Tong, W.; Shi, L. M.; Blair, R.; Perkins, R.; Branham, W.; Hass, B. S.; Xie, Q.; Dial, S. L.; Moland, C. L.; Sheehan, D. M. Structure−Activity Relationships for a Large Diverse Set of Natural, Synthetic, and Environmental Estrogens. *Chem. Res. Toxicol.* **2001**, *14*, 280−294.
(60) Hashimoto, Y. Structural Development of Synthetic Retinoids and Thalidomide-Related Molecules. *Cancer Chemother. Pharmacol.* **2003**, *52*, S12−S23.
(61) Douguet, D.; Thoreau, E.; Grassy, G. Quantitative Structure−Activity Relationship Studies of RAR α, *β*, *γ* Retinoid Agonists. *Quant. Struct.−Act. Relat.* **1999**, *18*, 107−123.
(62) Vuligonda, V.; Garst, M. E.; Chandraratna, R. A. S. Stereoselective Synthesis and Receptor Activity of Conformationally Defined Retinoid X Receptor Selective Ligands. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 589−594.
(63) Ohta, K.; Tsuji, M.; Kawachi, E.; Fukasawa, H.; Hashimoto, Y.; Shudo, K.; Kagechika, H. Potent Retinoid Synergists with a Diphenylamine Skeleton. *Biol. Pharm. Bull.* **1998**, *21*, 544−546.
(64) *Targeted Libraries Generation 1*; Enamine: Kiev, Ukraine; www.enamine.net.
(65) Saurat, J.-H. Retinoids and Psoriasis: Novel Issues in Retinoid Pharmacology and Implications for Psoriasis Treatment. *J. Am. Acad. Dermatol.* **1999**, *41*, S2−S6.
(66) Xu, Y.; Mayhugh, D.; Saeed, A.; Wang, X.; Thompson, R. C.; Dominianni, S. J.; Kauffman, R. F.; Singh, J.; Bean, J. S.; Bensch, W. R.; Barr, R. J.; Osborne, J.; Montrose-Rafizadeh, C.; Zink, R. W.; Yumibe, N. P.; Huang, N.; Luffer-Atlas, D.; Rungta, D.; Maise, D. E.; Mantlo, N. B. Design and Synthesis of a Potent and Selective Triazole-Based Peroxisome Proliferator-Activated Receptor α Agonist *J. Med. Chem.* **2003**, *46*, 5121−5124.
(67) Kurtz, T. W. Treating the Metabolic Syndrome: Telmisartan as a Peroxisome Proliferator-Activated Receptor *γ* Activator. *Acta Diabetol.* **2005**, *42*, S9−S16.
(68) Zhi, L.; Tegley, C. M.; Kallel, E. A.; Marschke, K. B.; Mais, D. E.; Gottardis, M. M.; Jones, T. K. 5-Aryl-1,2-Dihydrochromeno[3,4-f] Quinolines: A Novel Class of Non-Steroidal Human Progesterone Receptor Agonists. *J. Med. Chem.* **1998**, *41*, 291−302.
(69) Brandes, L. J.; MacDonald, L. M.; Bogdanovic, R. P. Evidence that the Antiestrogen Binding Site is a Histamine or Histamine-like Receptor. *Biochem. Biophys. Res. Comm.* **1985**, *126*, 905−910.
(70) Dhar, A.; Liu, S.; Klucik, J.; Berlin, K. D.; Madler, M. M.; Lu, S.; Ivey, R. T.; Zacheis, D.; Brown, C. W.; Nelson, E. C.; Birckbichler, P. J.; Benbrook, D. M. Synthesis, Structure−Activity Relationships, and RAR *γ*−Ligand Interactions of Nitrogen Heteroarotinoids. *J. Med. Chem.* **1999**, *42*, 3602−3614.
(71) Chandraratna, R. A. Tazarotene: First of a New Generation of Receptor-Selective Retinoids. *Br. J. Dermatol.* **1996**, *135*, 18−25.
(72) Hashimoto, Y.; Miyachi, H. Nuclear Receptor Antagonists Designed Based on the Helix-Folding Inhibition Hypothesis. *Bioorg. Med. Chem.* **2005**, *13*, 5080−5093.
(73) Ahmadiani, A.; Mandgary, A.; Sayyah, M. Anticonvulsant Effect of Flutamide on Seizures Induced by Pentylenetetrazole: Involvement of Benzodiazepine Receptors. *Epilepsia* **2003**, *44*, 629−635.
(74) Zhang, Q.; Muegge, I. Scaffold Hopping Through Virtual Screening Using 2D and 3D Similarity Descriptors: Ranking, Voting, and Consensus Scoring. *J. Med. Chem.* **2006**, *49*, 1536−1548.
(75) Mestres J. Representativity of Target Families in the Protein Data Bank: Impact for Family-Directed Structure-Based Drug Discovery. *Drug Discovery Today* **2005**, *10*, 1629−1637.

CI600300K