

# Increasing The Odds Of Hit Identification By Screening Against Receptor Homologs

Yuzong Chen, Congzhong Cai, Zerong Li, Lianyi Han and Jifeng Wang

**Abstract**— Increasing the odds of hit identification in screening is of significance for drug discovery. The odds for finding a hit are closely related either to the diversity of libraries or to the availability of focused libraries. There are no truly diverse libraries and it is difficult to design focused libraries without sufficient information. Hence it is helpful to consider alternative approaches that can enhance the odds using existing libraries. Multiple members of a protein family have been considered collectively in inhibitor design, on the basis of the correlation between protein families and ligands derived from specific compound classes. Such a correlation has been exploited in various drug discovery studies and a general receptor-homolog-based screening scheme may be devised. The feasibility of such a scheme in enhancing the odds of hit identification is discussed.

**Index Terms**—Homology, inhibitors, screening.

## I. INTRODUCTION

High-throughput screening and virtual screening has been extensively used in drug discovery [1, 2]. The odds for finding a hit depends on the diversity of compound libraries used [3]. There is no truly universal set of representative compounds and the screening has practically been conducted against subsets of molecules [4, 5x], which may result in useful hits being missed. This problem was illustrated recently by Oldenburg [3] in an example of two highly similar compounds of the steroid family, testosterone and estrogen, which differ only by a methyl group and a few double bonds. If only the first is included in a library for screening against estrogen receptor, estrogen would not be discovered. The same is true if one starts with estrogen and tests it against testosterone receptor.

Such a problem is likely of particular concern to screening for agonist/activator drugs that generally require more specific structural binding configuration than that of antagonist/inhibitor drugs. A search of Medline shows that less than 12% of the publications in drug screening are related to agonist/activator drugs, which may be partly due to the difficulty in finding an agonist or activator hit. Agonist/activator drugs constitute an important drug class. A search of the therapeutic target database [5] finds 44 targets of agonist/activator drugs, many of which are important receptors. Hence, methods for improving the odds of screening of agonist/activator drugs as well as

antagonist/inhibitor drugs are potentially useful in new drug discovery.

Increasing the size of libraries in a random fashion may not always be effective or practical for solving this problem [4]. Drug-like compounds have been found sparsely distributed through chemistry space [6]. As a result, the design of focused libraries can be a difficult task without sufficient information. Therefore it is helpful to consider alternative approaches that can potentially enhance the odds of hit identification without relying solely on the simple expansion of existing libraries. Recent developments in exploiting the correlation between protein families and ligands from specific compound classes point to a receptor-homolog-based screening scheme for improving the odds of hit identification.

## II. EXPLOITATION OF THE CORRELATION BETWEEN PROTEIN FAMILIES AND LIGANDS DERIVED FROM SPECIFIC COMPOUND CLASSES

In the search of inhibitors of specific cyclin-dependent kinases, multiple members of the kinase family were considered in a collective manner on the basis of their common feature of ligand-binding mode [7]. Based on the known binding mode of purine olomoucine at the ATP-binding site, compounds were designed from combinatorial libraries of 2,6,9-trisubstituted purines. Selective inhibitors for subsets of cyclin-dependent kinases were developed from these libraries.

The correlation between members of kinase family and inhibitors derived from specific compound classes has been shown and exploited in various studies [8]. For instance, compounds based on quinazoline scaffold were found to exhibit good structure-activity relationship against EGFR tyrosine kinase and other related kinases (30-34). Potent ATP-binding site inhibitors were derived from this scaffold for EGFR tyrosine kinase (31 or 32), c-erbB2/c-erbB4/EGFR (42, 45, 62), RAF kinase (61), CSF-1R (47, 51), and VEGFR (63-72), some of which are undergoing clinical trials. Inhibitors were derived from the phenylamino-pyrimidine class for PDGFR (80?), PKC- $\alpha$  (75) and EGFR tyrosine kinase (82). A number of indolocarbazole derivatives were found to be inhibitors of NGF receptor (210), protein kinase C (215), and PDGR (213). Pyrazolo[d]pyrimidine derivatives were designed as inhibitors of LcK (120), v-Src (220), CSF-1R (47), and EGFR (47).

This correlation also appears in other protein families and applies to agonists/activators. A few examples are serine protease and peptide-like inhibitors (x1), nuclear hormone receptors and steroid agonists (x2), and members of G-protein coupled receptors and catecholamine agonists (x2).

There have been suggestions of exploiting this relationship to various drug discovery problems [10]. It was proposed that, by screening a common and diverse set of small molecule inhibitors against a set of proteins from a family, specific structure-activity relationship homology can be derived from which potential drug discovery targets can be grouped (z1). Based on the common molecular theme for ligands with a certain classes of drug targets it was suggested that, in stead of putting barriers of high-risk targets through expensive screens of large compound collections, focused libraries of specific compound classes should be tested first [9]. In order to direct discovery processes to tractable chemical libraries, potential targets can be screened from all members of a gene family that have proven records in drug development and chemistry effort can then be focused on the most intriguing targets (z2).

#### *A. Screening against of receptor-homologs as a general scheme?*

It is of interest to explore the possibility of further extending the protein-family-based approach into a more general receptor-homolog-based scheme for drug screening. In this scheme, as illustrated in Figure 1, screening is conducted against a receptor and its homologs which are defined as proteins of similar sequence in the ligand-binding domain. These homologs likely share common structural features at ligand-binding sites and structurally similar ligands may exist for some of these homologs. If one or more of these ligands is identified as a hit for the entire homolog group, the rest may be generated by focused library design based on the identified hits. The subsequent screening of these focused libraries against the receptor may lead to the identification of the specific ligands for that receptor.

In the testosterone-estrogen example, if the same set of compounds is used for screening against estrogen receptor and all of its sequence homologs (including testosterone receptor), testosterone would be identified as a hit for the entire receptor group. A focused library of steroid analogs may be constructed based on the framework of testosterone, which likely include estrogen. A subsequent screening of this focused library against estrogen receptor would discover estrogen. The odds of finding an agonist hit are thus significantly increased.

The feasibility of this scheme depends on the existence of structurally similar ligands, particularly agonists/activators, for at least some of the homologs. Structurally similar agonists are searched for two receptors, estrogen receptor and adrenoceptor, and their sequence homologs. The identified homologs of these receptors along with structurally similar agonists are given in Table 1 and 2 respectively. Certain homologs are excluded from these two groups which include orphan receptors with unknown ligand and peptide-agonist

receptors or lipid-agonist receptors with binding sites covering sections different from that of the respective common binding sites defined by the estrogen receptor or adrenoceptor [11,12]. Structurally similar agonists are found for 7 out of 8 homologs of estrogen receptor [13-15] and for 4 out of 8 homologs of adrenoceptor [16-19] respectively. Moreover, agonist superficially similar to the other agonists is found for a homolog of estrogen receptor and that of adrenoceptor respectively. Thus it appears that compounds of common structural framework can be found for a substantial portion of receptor homologs.

### III. CONCLUSION

The correlation between protein families and ligands of common structural theme has been exploited in various drug discovery studies. A general receptor-homolog-based screening scheme may be devised based on this correlation. From the study of specific cases, it appears to be feasible to use this scheme to enhance the odds of hit identification.

#### References

- [1] Ohlstein, E. H. et al. (2000) Drug discovery in the next millennium. *Annu. Rev. Pharmacol. Toxicol.* 40, 177-191
- [2] Schneider, G. and Bohm, H-J. (2002) Virtual screening and fast automated docking methods. *Drug Discov. Today* 7, 64-70
- [3] Oldenburg, K.R. (2001) Optimizing screening technology: how much to invest? - Reply. *Drug Discov. Today* 6, 128-129
- [4] Hann, M. and Green, R. (1999) Chemoinformatics - a new name for an old problem? *Curr. Opin. Chem. Biol.* 3, 379-383
- [4] Chen, X. et al. (2002) TTD: Therapeutic Target Database. *Nucleic Acids Res.*30, 412-415
- [5] Lipinski, C.A. (2000) Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods.* 44, 235-249
- [6] Gray, N. S. et al. (1998) Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* 281, 533-537
- [7] Garcia\_Echeverria, C. et al. (2000) ATP site-directed competitive and irreversible inhibitors of protein kinases. *Med. Res. Rev.* 20, 28-57
- [8] Thorpe, D.S. (2000) Forecasting roles of combinatorial chemistry in the age of genomically derived drug discovery targets. *Comb. Chem. High Throughput Screen.* 3, 421-436
- [9] Caron, P. R. et al. (2001) Chemogenomic approaches to drug discovery. *Curr. Opin. Chem. Biol.* 5, 464-470
- [10] Costantino, G. and Pellicciari, R. (1996) Homology modeling of metabotropic glutamate receptors. *Structural*

motifs affecting binding modes and pharmacological profile of mGluR1 agonists and competitive antagonists. *J. Med. Chem.* 39, 3998-4006

[11] Hawtin, S.R. and Wheatley, M. (2000) Critical role of a subdomain of the N-terminus of the V1a vasopressin receptor for binding agonists but not antagonists; Functional rescue by the oxytocin receptor N-terminus. *Biochemistry* 39, 13524-13533

[12] Steinmetz, A.C.U. et al. (2001) Binding of ligands and activation of transcription by nuclear receptors. *Annu. Rev. Biophys. Biomol. Struct.* 30, 329-359

[13] Lala, D.S. et al. (1997) Activation of the orphan nuclear receptor steroidogenic factor-1 by oxysteroids. *Proc. Natl. Acad. Sci. USA* 94, 4895-4900

[14] Weatherman, R.V. et al. (1999) Nuclear-receptor ligands and ligand-binding domains. *Annu. Rev. Biochem.* 68, 559-581

[15] Rang, H.P. et al. (1999) *Pharmacology* (4th edn) (Hunter, L. ed) Churchill livingstone

[16] Arvidsson, L. E. et al. (1984) 8-Hydroxy-2-(alkylamino) tetralins and related compounds as central 5-hydroxytryptamine receptor agonists. *J. Med. Chem.* 27,45-51

[17] Van der Goot, H. and Timmerman, H. (2000) Selective ligands as tools to study histamine receptors. *J. Med. Chem.* 35, 5-20

[18] Gerhardt, C. C. et al. (1997) Molecular cloning and pharmacological characterization of a molluscan octopamine receptor. *Molecular Pharmacology.* 51, 293-300