eISSN: 09748369, www.biolmedonline.com

Indispensable chemical genomic approaches in novel systemic targeted drug discovery

A Kanikkai Raja, A Babu Vimalanathan, S Vinoth Raj, S Suresh Kumar, G Swamynathan, *Manoj G Tyagi

Department of Pharmacology, Christian Medical College, Vellore 632002, Tamilnadu, India.

***Corresponding Author:** tyagi239@yahoo.co.in

Abstract

Chemogenomics is a new emerging area in the field of drug discovery and development. It describes the development of target specific chemical ligands and the use of such specific chemical ligands to globally study the gene level and protein level functions. The human genome contains around 100,000 genes and 30,000 proteins which are encoded by it. Primarily the novel, systemic, very small molecular sized, cell permeable and target specific chemical ligands are particularly useful in systematic genomic approaches to study the normal and abnormal biological functions. The complete genomic sequence information with structural and comparative genomics when combined with all the features in synthetic chemistry, ligand screening and identification provides target or functions specific chemical ligands and drugs. Currently, the in-silico approaches are in vogue in novel target prediction and systemic drug discovery. It is prediction of biological targets of small molecules via data mining in target annotated chemical databases. This review will focus on the chemogenomics, its approaches for rational drug design and the recent in-silico approaches in novel target prediction.

Keywords: Drugs; Chemotherapeutic; Chemogenomics; Human genome; Genomic sequence; Chemical ligands; Structural and comparative genomics; Synthetic chemistry; In silico.

Introduction

The completion of the human genome suggests that there are tens of thousands of genes (Lander *et al.*, 2001; Ventoer *et al*., 2001) and at least as many proteins, many of these proteins are potential targets for drug intervention to control human disease or injury. It has been estimated that, out of 20-25,000 human genes supposed to encode for ca. 3000 drugable targets (Russ and Lampel, 2005), only a subset of that pharmacological space (ca. of 800 proteins) has currently been investigated by the pharmaceutical industry (Drews and Ryser, 1997). Progress in meeting this challenge will contribute to further the fundamental understanding of the biological function of the individual gene products and ultimately provide a basis for the discovery of new and better therapies for disease. Chemogenomics, chemical genomics chemical genetics and chemical biology are the new front running drug-discovery technologies that aim to address this scientific challenge and that are not only semantically related. The chemical genomics/chemo genomics is defined as the systemic identification of small molecules that interact via a specific molecular recognition mode with target proteins encoded by the genome. The term Chemogenomics is applied more specifically to target family approaches in drug discovery. Chemical genetics is an analogous way of identification of genetic mutations, to identify chemical compounds which induce or revert specific biological phenotypes by using cell –based or microorganism-based screening of compounds. Chemical biology is defined as the functional and mechanistic investigation of biological systems using chemical compounds and constitutes a more general discipline. They integrate basic disciplines like chemistry, genetics, chemo and bioinformatics, structural biology and biological screening into phenotypic and target based assays (Bradel and Jacoby, 2004; Mannhold *et al* 2004; Stockwell 2000; Schreiber, 2005).

Efficient target prediction

The chemogenomics terminology itself originated from the directed exploration of target gene families which was probably first emphasized in 1996 by researchers at Glaxo Welcome, who discussed the concept of systematization of drug discovery within target families like G-protein coupled receptors, ion channels or protease that had hitherto been successfully explored. The Glaxo Welcome scientists highlighted obvious advantages of

system-based approaches, such as combining advances in gene cloning and expression, automation, combinatorial chemistry and bioinformatics (Lehmann *et al* 1996). Directly following from the lock and key analogy for small molecules and their receptors in the human body, the question arises, into which locks will a key fit, i.e., with which targets will a compound bind. Early medicines such as Aspirin and its precursors have been used for thousands of years without knowing the protein targets, and their relatively safety has been established on use empirically. This situation also applies to some current drugs on the market such as the Bcr-Abl/c-KIT inhibitor, Imatinib, where interestingly promiscuity has been increasingly viewed in a positive light because it attacks at multiple points in signaling pathways which may be more efficient than antagonizing single targets and further, may level the risk of developing drug resistance (Hampton, 2004). Still in order to conduct proper clinical trials (both ethically and economically), it is beneficial to know as much as possible about the system, one is investigating. Both desired on-target effect as well as undesired off-target effects can be better understood by knowing the targets modulated by a given drug. This is beneficial on both accounts: for the desired targets, establishing the mode of action of active ingredient and facilitating its rational optimization (e.g. crystallization of the target protein). In addition off-targets completely unrelated to the actual target can be established enabling us to predict other secondary side effects (Jenkins *et al* 2007).

Before the advent of molecular biology and protein expression systems, all drug discoveries were necessarily systems based drug discovery. However, over the past few decades a shift towards target-based screening took place, leading to higher-throughput biochemical assays. This step towards target-based screening was caused by a selectivity paradigm which maintained that selective interference with metabolic or signaling networks is an ideal characteristic for medicines, thus promoting drugs to be as selective as possible. More recently it was realized that cellular networks are often able to compensate for single-point modification (Csermely *et al* 2005), resulting in less effective values. Several experimental strategies for elucidating the target of a particular compound exist such as affinity chromatography, expression cloning and protein microarrays (Hart 2005). Still, these approaches usually require a large amount of time and capital expenditure. Recently, *in silico* concepts have been used to predict particular compound's mode-of-action. Firstly, statistical models can be generated based on the structure of compounds known to show a certain degree of bioactivity, which is undoubtedly one of the major applications of the chemogenomics database. While structure based chemical descriptors have been proven to work quite well for target prediction (Nidhi *et al*, 2006) in recent years a shift from structure derived descriptors to biological descriptors could be observed. The affinity of a compound to a panel of targets was used to generate models of a very diverse nature. For example, affinity fingerprints can be used to make predictions about active targets, about adverse reactions or about pathways modulated by a particular compound (Kauvar *et al*, 1995) (Fig:1).

***In Silico* approaches in target prediction**

Chemical similarity searching (Bender and Glen, 2004) for target prediction is composed of a compound structure from a database of compounds with known targets. By employing the molecular similarity principle, the target of similar compounds may also be the target of the query structure. Likewise, it is (within limitation) well established that similar targets bind to similar compounds (Schuffenhauer *et al*, 2003). Thus, conventional similarity searching is inverted usually when new ligands for a known target are desired, in this case new targets for a known ligand are proposed. Most commonly due to their speed, 2D descriptors such as circular finger prints (Glen *et al* 2006) are employed as target prediction approaches (Nidhi *et al*, 2006). Similarity searching as a method for target fishing (i.e. identification) can also be performed with 3D chemical descriptors (Nettles *et al* 2006). In this case it is suggested that while 2D descriptors are powerful for similarity searching in annotated databases, 3D descriptors are often more appropriate when the orphan compound has low 2D similarity to all database and molecules (Nettles *et al* 2006).

Data miming in annotated chemical database is second somewhat more sophisticated method of predicting targets for small molecules. Multi-dimensional models resulting from data miming differ from similarity searching in that information from multiple ligands can be considered in parallel to make target suggestions. One caveat however that is, a systemic nomenclature for target information is

required for forming distinct activity class sets. As one of the first such applications, the PASS (prediction of activity spectra for substances) application used circular finger prints to train a Bayesian type classifier on the target activities listed in the MDDR database (Poroikov *et al* 2000) An approach based on a more comprehensive list of targets has been presented more recently, using circular finger prints and a Bayesian classifier but this time trained on a total of 964 targets (Nidhi *et al*, 2006).

Combinatorial and de novo approaches in chemical ligand design

Combination organic synthesis has increasingly been used to generate small molecule libraries for ligand screening and drug discovery. It stems from the original solid –phase oligonucleotide and peptide synthesis which provides a way for synthesis and purity of peptides with defined sequences (Blocker, 1983; Geysen *et al* 1984). Such technique also allows synthesis of oligonucleotides or peptides of all possible permutation of nucleotides or amino acids residues. This concept has been extended to other small molecules as well. Polycyclic compounds and natural product derivatives that mimic biological active molecules (biomimetics) are one among the most common types of structures that have been made in this way (Schreiber, 2000; Gray, 2001). This is because small molecule libraries based on a known lead compound (also known as focused libraries) often provide better chances in finding new biologically active molecules.

Another concept in chemical ligand synthesis that has emerged in recent years is to change the chirality of the compounds. Biological molecules often contain chiral or asymmetric carbon atoms, at which the carbon atom is bonded to four different moieties. Such molecules exist in two forms, known as enantiomers. Theses molecules have opposite effect in rotating plane-polarized light and their three dimensional structure are different and therefore possess different biological properties. The best example is the twenty naturally occurring amino acids. Metabolism (enzymes) of a living organism is specific for particular enantiomers such as yeast can only metabolize L-amino acids, which rotate plane-polarized light to the left, but not D-amino acids, which rotate plane-polarized light to the right (Hutt and Grady, 1996; Vuylsteke, 2000).

Lead compounds (*de novo* design) are provided by these methods to develop a meaningful chemical library that can eventually yield biologically active compounds. To develop successful lead compounds against diverse protein targets, and a designer approach is needed. This is where the *in silico* approach comes into play, owing to significant advances in computational power and the development of artificial intelligence (AI) and database that details known molecular structures with their corresponding functions from basis of the so–called Quantitative Structure Activity Relationship (QSAR) approaches. Fig: 2 indicates the structure–activity relationship. QSAR and its variants provide the statistical frame work for analysis of virtual libraries of small molecules (Klebe, 2000). Computer programs and algorithms have been developed to guide every major step in *de novo* design. Programs such as GASP (Generation Automatique de Structures Polycycliques) which generate a virtual library of large number of polycyclic structures with given parameters, can help to expand the diversity and complexity of potential chemical ligands (Barone *et al* 2001). Another area worthy of attention is the use of Darwinian molecular evolution in ligand design. It dates back to 1990 when Tuerk and Gold looked for RNA ligands that bind to bacteriophage T4 DNA polymerase by using Systematic Evolution of Ligands by Exponential Enrichment (SELE X) (Tuerk and Gold, 1990).

Comparative and structural genomics in systemic chemical ligand design

Comparison of whole genome sequences of different organisms is a powerful tool for identifying essential components of cellular processes, understanding the genome architecture as well as reconstructing the evolutionary events. Tatusov *et al* (1997) reported the comparison of 7 complete genomes in 5 major phylogenetic lineages and this has now expanded to 44 complete genomes in 30 major phylogenetic lineages. The database of Clusters of Orthologous Groups (COGs) allows information flow from well-studied model organisms to poorly characterized ones. In recent years, comparative genomics in mammals has improved the resolution of genetics maps in human and other model organisms and identified many human disease genes in other mammalian models (O'Brien, 1999).

Information obtained from comparative genomics studies is having a significant impact on the field of structural biology. It is clear that the three dimensional (3D) structures of proteins can greatly facilitate the drug discovery processes. In many cases, the availability of crystal structure data or predicted structure not only helps us to better understand the catalytic activity of the protein, but also shows the mechanism of how such a protein interact with other proteins as well as the allosteric actions or conformational changes involved. The continuous improvement in computation power and development of new bioinformatics tools have laid foundation for high-throughput X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy in protein structure determination (Stevens, 2000; Stevens *et al* 2001). These structural genomic approaches are aimed at systematic generation of experimental protein structures and unique structural folds (Burley, 2000). In order to provide structures for each family of proteins using these high-throughput analysis techniques, efforts have been made to ensure proper protein target selection (Brenner, 2000; Ramya *et al* 2007).

High throughput or high content screening technologies in drug discovery

Chemical ligand discovery through natural product screening has spearheaded chemotherapeutic development and biological discovery for many years (Bull *et al* 2000; Stockwell *et al*, 1999). Synthetic chemistry is now capable of producing large numbers of diversified chemical compounds which requires high-throughput screening (HTS) technologies for selection of specific ligands. One prominent example of HTS is In vitro Cell Line Screening Project (IVCLSP) under the Developmental Therapeutic Program (DTP) at National Cancer Institute (NCI), which uses 60 human cancer cell line (NCI 60) for the drug discovery screen. About 5,000 out of 60,000 potential anticancer compounds screened are found to have anti-cancer activity. After eliminating closely related compounds and the ones with known mechanisms of action, the remaining 1,200 were selected for further testing. Another example of cell based HTS is the Receptor Selection and Amplification Technology (R-SAT), which produces a signal (usually in terms of changes in the levels of β-galactosidase) whenever a test compound interacts with an expressed receptors such as G-protein coupled receptors and nuclear receptors. In the presence of receptors' activity, cells overcome contact inhibition and proliferate. The increase in cell number is translated into an increase in the level of β-galactosidase, whose level can be quantitatively measured using a spectrophotometer.

Stockwell *et al* (2000) developed a miniaturized, whole cell immunodetection system, called "cytoblot" for high throughput screening of biologically active molecules. Like DNA, small chemical ligands can also be immobilized on a glass surface for *in vitro* ligand screening (McBeath *et al*, 1999). Knocaert *et al* reported the use of such an approach in identifying intracellular targets of chemical inhibitors of the cyclin dependent protein kinase (Knocaert *et al*, 2000). Several groups have recently reported the development of protein micro-arrays for variety of applications such as identification of substrates for different protein kinases (Zhu *et al* 2000), antibody screening (Lueking *et al* 1999), screening for protein-protein interactions as well as protein targets for chemical probes (McBeath *et al*, 2000) and accurate measurement of protein abundance using antibody/antigen microarrays (Haab *et al* 2001).

More recently, McDonald *et al* (2006) reported on a high-content assay where they monitored signal transduction pathways by detecting protein-protein interactions within the spatial cellular context, described as protein complementation assay (PCA) (Michnick *et al* 2000). These protein complementation assays specifically use the cells which are engineered to simultaneously express two proteins from a common signaling pathway that are fused to complementary fragment of a fluorescent protein receptor. When the two proteins physically interact, they bring together the complimentary fragments and generate a fluorescent report. The group constructed 49 such PCAs that probed a diverse set of cellular pathways (for example, cell cycle, ubiquitin proteolytic and stress response), and monitored the activity and intracellular localization of those pathways independently by automated microscopy response to 107 different bioactive compounds. Hierarchical clustering of compounds based on the pathway activity biological fingerprints successfully recapitulated known structure activity relationships.

Chemical ligands and gene expression profiling

The most popular approach currently is to use chemical ligands to study the role of drug target

in the control of gene expression. High-density oligonucleotide or complementary DNA (cDNA) microarrays (collectively called DNA microarrays) immobilized on the surface of glass slide can be used simultaneously to analyze the expression of thousands of genes (Lockhart and Winzeler, 2000). DNA microarrays have proven to be very useful in elucidating the functions of the drug target involved in transcriptional control (Fig: 3). In such an approach, the cell, tissue or organism is treated with chemical ligands. mRNA are then prepared from the treated and untreated cell or organism, and then used to produce fluorescence labeled cDNA to hybridizes DNA microarrays and generate gene expression profiles. By comparing the differences between the profiles before and after drug treatment, genes whose expression is modulated by the chemical ligands are then identified. These genes are further classified into pools based on similar functions or co-regulation in the cell. This information can often reveal specific transcription factors and other regulators for each gene pool, thereby allowing assembly of potential regulatory pathways involving the drug target. The good example is provided by a recent study of yeast histone deacedylases (HDACs) (Bernstein *et al* 2000).

Chemical ligand and its global genetic interaction

Genetic interaction is powerful way to study the relationship among genes or protein. If a gene acts in the same pathway or a parallel pathway with a second gene, its mutation may affect the phenotype(s) of a mutation in the second gene. Such secondary mutations lead to modification of phenotypes of the first mutations and are thus called modifier mutations, which allows the identification of new components in the pathway(s) of a gene of interest. In extreme cases, two normally viable mutations, when combined, can generate a lethal phenotype, which is called synthetic lethality. Several genomic deletion projects have systematically generated deletion of individual genes in model organism including the budding yeast (Anderson *et al* 1999), the worms (Liu *et al* 1999; Bergmann, 2001), the fruit fly (Spradling *et al* 1999) and the mouse (Coelh *et al* 2000).

The potential global genetic pathways or global genetic network can then be assembled by simply pooling genes with similar cellular functions (Fig 4). The screening process can be easily automated either as a 96-well plate-based assay or as an oligonucleotide microarray-based assay that takes advantage of the internal barcodes to each deleted locus (Giaever *et al* 1999). Giaever *et al.* demonstrated the feasibility of a genome wide drug sensitivity screen by examining 233 yeast deletion strains to the drug tunicamycin. The yeast genomic deletion mutants were first used in the global study of genetic interactions with TOR (Chan *et al* 2000). TOR is a highly conserved ataxia telangiectasia–related protein kinase essential for cell growth. To establish a global genetic interaction network of TOR, Chan *et al* systematically measured the sensitivity of individual yeast mutants to a low concentration of rapamycin based on the relative growth of each mutant, thereby assembling a global genetic interaction network for TOR. The genome-wide screen is obviously advantageous, because it profiles every single mutant gene, regardless of the severity of the mutants' phenotype. Even a moderate sensitivity could be biologically significant, which would likely be missed in a traditional genetic screen due to a weak phenotype. Moreover, every single mutation is a complete deletion - it avoids many complicated phenotypes by a point mutation or multi-allelic mutations in a typical traditional genetic screen.

Chemogenomic examples

An example of a broad family of enzymes, many of which are attractive drug targets because of their roles in signal transduction, is the protein kinase (s). Roughly, 500 serine/threonine and tyrosine protein kinase have homologous sequences. This sequence homology is also evident in the structural similarity of proteins within the protein kinase family. Inhibitors such as staurosporine are active against both serine/threonine and tyrosine kinase and can serve as a starting point for discovery of highly specific inhibitors (Garcia-Echeverria *et al*; Cohen, 1999). Relatively simple modifications of staurosporine have been shown to increase specificity consistent with predictions from homology modeling of cross-reacting kinase (Lamers *et al* 1999). Recent interest in the mitogen activated protein (MAP) kinase have led to the discovery of multiple parallel protein kinase cascades comprised of proteins of related sequences (Treisman, 1996). Another example of an interesting gene family is the caspase, which are cysteine proteases with specificity for cleavage after aspartyl residues. Interleukin-1ß converting enzyme (ICE) has been shown to be essential for cytokine

processing and is currently being pursued as a drug target (Dinarello, 1998).There are also roughly a dozen caspases with sequence homology to ICE, and although the exact function of all of these is not known, it is clear that some of these have an important role in regulation of apoptosis (Marks and Breg,1999). Structural insights through X-ray crystallography facilitated the rapid identification of selective inhibitors of these other potential drug targets. Experience with ICE has also been applied to the other caspase-family members in the areas of expression, purification, assay development, crystallization and structure determination (Wei *et al* 2000).

Conclusion

Chemogenomics approaches to rational drug discovery have been under investigation in the last years as high-throughput data became available for both targets and ligands of pharmaceutical interest. Real power of having the complete genome sequence lies with the ability to use this information to develop novel therapeutics. The chemogenomics approaches, in which this sequence information is combined with protein structure and models to link to chemical inhibitors or activators, is designed to fully utilize the sequence information by considering the large families of gene targets at once, which depends on various well established approaches, all of which drive the efficient utilization of information and capital. Chemogenomics approaches can produce complete biological and genetic profiles of the drug target protein, examination of which will allow accurate detailed prediction and assembly of the biological pathways involved. With chemogenomics approaches, bottlenecks of the drug development has been shifted to the target-validation area but it also provides reagents that, if properly designed, can help alleviate this new challenge by taking advantage of availability of compounds such as biochemical probes to dissect the role of novel target at level of cell or animal-disease model. This knowledge will also be helpful for understanding disease mechanism and selecting the most relevant protein targets for discovery of the optional drugs, predicting other benefits and side effects of the existing drugs based on the global effect of the drug on gene expression, posttranslational modification and the biological pathways based on various chemogenomics approaches. In near future, an efficient, easier and better control of ligand selectivity can be ensured by using these data and novel genetic target can be better addressed in the target space.

Acknowledgement

We are thankful to Shri D Kesavan for his help in the preparation of figures.

References

Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, MChu A, Connelly C, Davis K, Dietrich F, Dow SW, ElBakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Layub M, Laub H, Liao H, Davis RW, 1999 . Functional characterization of the S.cerevisiae genome by gene deletion and parallel analysis. Science, 285:901-6.

Bargmann CI, 2001. High-throughput reverse genetics: RNAi screen in Caenorhabditis elegans. Genome Biology, 2: REVIEWS1005.1.

Bredel M, Jacoby E, 2004. Chemogenomics: an emerging strategy for rapid target and drug discovery. Nature Reviews Genetics, 5:262–275.

Bender A, Glen RC, 2004. Molecular similarity: a key technique in molecular informatics. Organic and Biomolecular Chemistry, 2:3204-3218.

Blocker, 1983. A new general approach for the simultaneous chemical synthesis of large numbers of oligonucleotides: segmental solid supports. Nucleic Acids Research, 11:4365-77.

Barone R, Arbelot M, Chanon M, 2001. GASP: a computer program to generate automatically polycyclic structures. Tetrahedron, 57:6035-6042.

Bull AT, Ward AC, Goodfellow M, 2000. Search and discovery strategies for biotechnology: the paradigm shift. Microbiology and Molecular Biology Reviews, 64:573-606.

Bernstein BE, Tong JK, Schreiber SL, 2000. Genome-wide studies of histone deacetylase function in yeast. Proceedings of National Academy of Sciences USA, 97:13708-13.

Burley SK, 2000. An overview of structural genomics. Nature Structural Biology, 7:932-4.

Brenner SE, 2000. Target selection for structural genomics. Nature Structural Biology, 7:967-9.

Csermely P, Agoston V, Pongor S, 2005. The efficiency of multi-target drugs: the

network approach might help drug design. Trends in Pharmacological Sciences, 26:178–182.

Coelh PS, Kumar A, Snyder M, 2000. Genome-wide mutant collections: toolboxes for functional genomics. Current Opinions in Microbiology, 3:309-15.

Chan TF, Carvalho J, Riles L, Zheng XFS, 2000. A chemical genomics approach towards understanding the global functions of TOR. Proceedings of National Academy of Sciences USA, 97:13227-13232.

Cohen P, 1999. The development and therapeutic potential of protein kinase inhibitors. Current Opinions in Chemical Biology, 3:459-465.

Drews J, Ryser S, 1997. The role of innovation in drug development. Nature Biotechnology, 15:1318-1319.

Dinarello CA, 1998. Interleukin-1 beta, Interleukin-18, and the Interleukin-1 beta converting enzyme. Annals of New York Academy of Sciences, 856:1-11.

Geysen HM, Meloen RH, Brateling SJ, 1984. Use of peptide synthesis to prove viral antigens for epitopes to a resolution of a single amino acids. Proceedings of National Academy of Sciences USA, 81(13):3998-4002.

Glen RC, Bender A, Amby CH, Carlsson L, Boyer S, Smith JI, 2006. Drugs, 9:199.

Gray NS, 2001. Combinatorial libraries and biological discovery. Current Opinions in Neurobiology, 11:608 14.

Giaever G, Soemarker DD, Jones TW, Liang H, Winzeler EA, Astromoff A, Davis RW, 1999. Genomic profiling of drug sensitivities via induced haploinsufficiency. Nature Genetics, 21:278-83.

Garcia-Echeverria C, Traxler P, Evans DB, 2000. ATP site-directed competitive and irreversible inhibitors of protein kinases. Medicinal Research Reviews, 20:28-57.

Hampton T, 2004. “Promiscuous" anticancer drugs that hit multiple targets may thwart resistance. JAMA, 292:419-422.

Hart CP, 2005. Finding the target after screening the phenotype. Drug Discovery Today, 10:513-519.

Hutt AJ, Grady JO, 1996. Drug Chirality: a consideration of the significance of the stereochemistry of antimicrobial agents. Journal of Antimicrobial Chemotherapy, 37:7-32.

Haab BB, Dunham MJ, Brown PO, 2001. Protein micro arrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. Genome Biology, 2:RESEARCH0004.

Jenkins JL, Bender A, Davies JW, 2007. *In silico* target fishing: Predicting biological targets from chemical structure. Drug Discovery Today Technology, 3:413-21.

Kauvar LM, Higgins DL, Villar HO, Sportsman JR, Engqvist-Goldstein A, Bukar R, Bauer KE, Dilley H, Rocke DM, 1995. Predicting ligand binding to proteins by affinity fingerprinting. Chemistry and Biology, 2:107-108.

Klebe G, 2000. Recent developments in structure-based drug design. Molecular Medicine, 78:269-81.

Knocaert M, Gray N, Damiens E, Chang YT, Grellier P, Grant K, Fergusson D, Mottram J, Soete M, Dubremetz JF, Le_Roch K, Doerig C, Schultz P, Meijier L, 2000. Intracellular targets of cyclin-dependent kinase inhibitors: identification by affinity chromatography using immobilised inhibitors. Chemistry and Biology, 7:411-22.

Lehmann J, Boxter A, Brown D, Connolly P, Geysin M, Hayes M, Howard R, Knowles J, Lee M, Lyall A, Maran N, Niebel J, Rees E, Saussy S, Short R, Tarbir M, Ward M, Weitkamp E, Williamson R, 1996. Nature, 384 SUPP:1-5.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Doyle M, Fitzhugh W, 2001. Initial sequencing and analysis of the human genome. Nature, 409:860-921.

Lamers MB, Antson AA, Hubbard RE, Sott RK, Willams DH, 1999. Structure of the protein tyrosine kinase domain of C-terminal Src kinase (CSK) in complex with staurosporine. Journal of Molecular Biology, 285:713-725.

Lueking A, Horn M, Eickhoff H, Bussow K, Lehrach H, Walter G, 1999. Protein micro arrays for gene expression and antibody screening. Annals of Biochemistry, 270:103-11.

Lockhart DJ, Winzeler EA, 2000. Genomics, gene expression and DNA arrays. Nature, 405:827-36.

Liu LX, Spoerke JM, Mulligan EL, Chen J, Reardon B, Westlund B, Sun L, Abel K, Armstrong B, Hardiman G, King J, McCaague L, Basson M, Clover R, Jhonson CD, 1999. High throughput isolation of Caenorhabditis elegans deletion mutants. Genome Research, 9:859-67.

Marks N, Breg MJ, 1999. Recent advances on neuronal caspases in development and neurodegeneration. Neurochemistry International, 35:195-220.

MacBeath G, Schreiber SL, 2000. Printing proteins as micro arrays for high-throughput function determination. Science, 289:1760-3.

MacBeath G, Koehler A, Schreiber S, 1999. Printing small molecules as micro arrays and detecting protein-ligand interactions *en masse*. Journal of American Chemical Society, 121:7967-7968.

MacDonald ML, Lameridin J, Owens S, Keon BH, Bilter GK, Shang Z, Huang Z, Yu H, Dias J, Minami T, Michnick SW, Westwick JK, 2006. Identifying off-target effects and hidden phenotypes of drugs in human cells. Nature Chemical Biology, 2:239.

Michnick SW, Remy I, Campbell-Valois FX, Vallee-Belisle A, Pelletier JN, 2000. Detection of protein-protein interactions by protein fragment complementation strategies. Methods in Enzymology, 328:208-30.

Nidhi, Glick M, Davies JW, Jenkins JL, 2006. Prediction of biological targets for compounds using multiple-category Bayesian models trained on chemogenomics databases. Journal of Chemical Information and Modeling, 46:1124-1133.

Nettles JH, Jenkins JL, Bender A, Deng Z, Davies JW, Glick M, 2006. Bridging chemical and biological space: "target fishing" using 2D and 3D molecular descriptors. Journal of Medicinal Chemistry, 49:6802-6810.

O'Brien SJ, Menotti-Raymond M, Murphy WJ, Nash WG, Wienberg J, Stanyon R, Copeland NG, Jenkins NA, Womack JE, Marshall Graves JA, 1999. The promise of comparative genomics in mammals. Science, 286:458-62,479-81.

Poroikov VV, Filimonov DA, Borodina YV, Lagunin AA, Kos A, 2000. Robustness of biological activity spectra predicting by computer program PASS for noncongeneric sets of chemical compounds. Journal of Chemical Information and Computer Sciences, 40:1349-1355.

Russ AP, Lampel S, 2005. The druggable genome: an update. Drug Discovery Today, 10: 1607-1610.

Mannhold R, Kubinyi H, Folkers G, 2004. Chomogenomics in Drug Discovery – A Medicinal Chemistry Perspective (Volume 22 of Methods and Principles in Medicinal Chemistry), ed. Wiley-VCH, Weinheim.

Ramya R, Chandran P, Prasanna CG, Gunaseelan S, Tyagi MG, 2007. Analysis of the effect of F287K mutation on the binding of vasopressin analog to the V_2 receptor. Indian Journal of Multiresolution, 3(1), 105-110.

Reed MA, Snyder M, 2000. Analysis of yeast protein kinase using protein chips. Nature Genetics, 26:283-9.

Stockwell BR, 2000. Chemical genetics: ligand-based discovery of gene function. Nature Reviews Genetics, 1:116-125.

Schreiber SL, 2005. Small molecules: the missing link in the central dogma. Nature Chemical Biology, 1(2): 64-66.

Schuffenhauer A, Floersheim P, Acklin P, Jacoby EJ, 2003. Similarity metrics for ligands reflecting the similarity of the target proteins. Chemical Information and Computer Sciences, 43:391.

Schreiber S, 2000. Target–oriented and diversity-oriented organic synthesis in drug discovery. Science, 287:1964-1969.

Stevens RC, 2000. High-throughput protein crystalisation. Current Opinions in Structural Biology, 10:558-63.

Stevens RC, Yokoyama S, Wilson IA, 2001. Global efforts in structural genomics. Science, 294:89-92.

Spradling AC, Stern D, Beaton A, Rhem EJ, Laverty T, Mozden N, Misra S, Rubin GM, 1999. The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. Genetics, 153:135-77.

Stockwell BR, Haggary SJ, Schreiber SL, 1999. High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modification. Chemistry and Biology, 6:71-83.

Tuerk C, Gold L, 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science, 249:505-10.

Tatuso RL, Koonin EV, Lipman DJ, 1997. A Genomic perspective on protein families. Science, 278:631-7.

Treisman R, 1996. Regulation of transcription by MAP kinase cascades. Current Opinions in Cell Biology, 8:205-215.

Vuylsteke A, 2000. Fundamentals of chirality. Acta Anaesthesiologica Belgica, 51:113-7.

Ventoer JC, Adams MD, Meyers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandall M, Evans CA, Holt RA, 2001. The sequence of the genome. Science, 291:1304-1351.

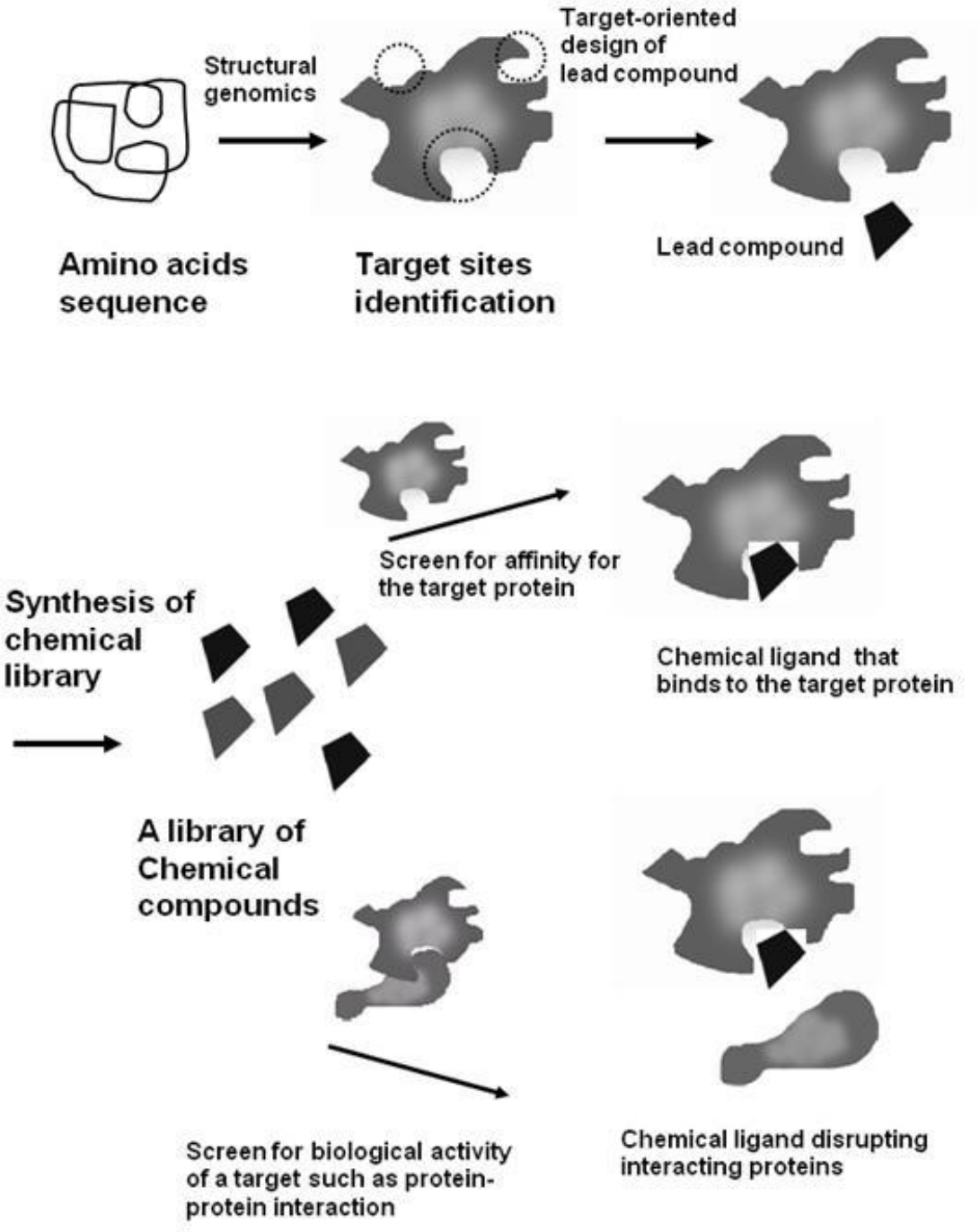
Wei Y, Fox T, Chambers SP, Sintchak J, Coll JT, Golec JM, Swenson L, Wilson KP, Charifson PS, 2000. The structures of caspase-1,-3,-7, and -8 reveal

the basis for substrate and inhibitor selectivity. Chemistry and Biology, 7:423-432.

Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M, 2000. Analysis of yeast protein kinase using protein chips. Nature Genetics, 26: 283-9.

Figures follow……

Figure 1 : A typical streamline of design, synthesis and selection of chemical ligands



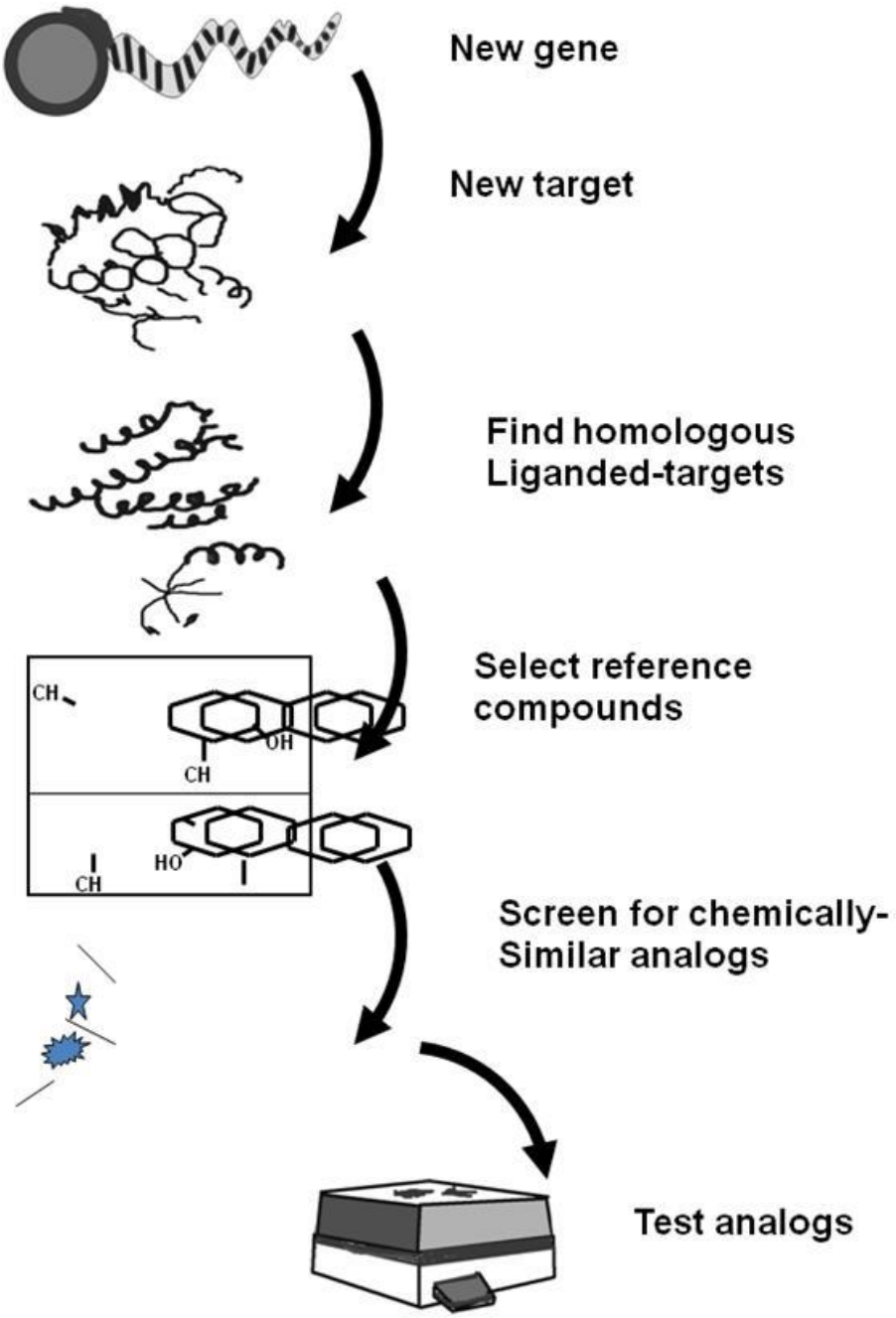


Figure 2 : Structure-activity relationship homology flowchart

Figure-3: Chemical ligands in gene expression and protein profilings

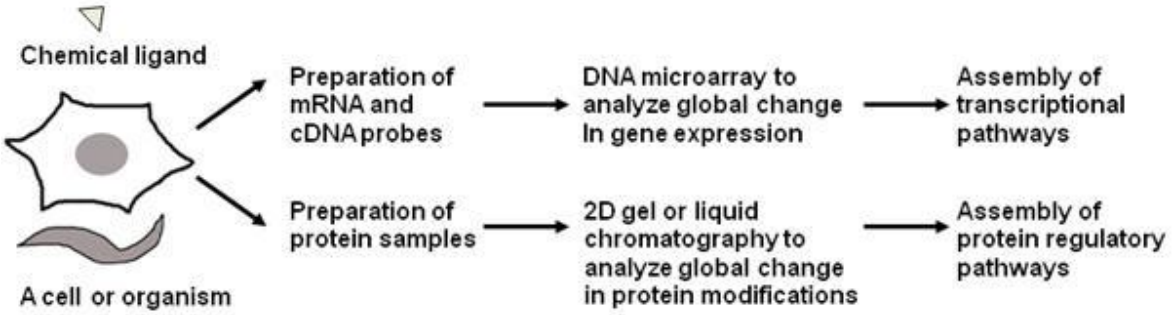


Figure-4: Chemical ligands in global genetic interaction study