Chemical genomics as an emerging paradigm for postgenomic drug discovery

F Raymond Salemme

President & CSO,
3-Dimensional
Pharmaceuticals, Inc.,
Three Lower Makefield
Corporate Center, 1020 Stony
Hill Road, Suite 300,
Yardley, PA 19067, USA
Tel: +1 267 757 7233;
Fax: +1 267 756 7276;
E-mail: salemme@3dp.com

Chemical genomics approaches are evolving to overcome key problems limiting the efficiency of drug discovery in the postgenomic era. Many of these stem from the low success rates in finding drugs for novel genomics targets whose biochemical properties and therapeutic relevance is poorly understood. The fundamental objective of chemical genomics is to find and optimize chemical compounds that can be used to directly test the therapeutic relevance of new targets revealed through genome sequencing. An integrated approach to chemical genomics encompasses a diverse set of tools including quantitative affinity-based screens, computer-directed combinatorial chemistry, and structure-based drug design. The approach is most effectively applied across targets classes whose members are structurally related, and where some members are known to have *bona fide* therapeutic relevance.

Keywords: chemical genomics, drug discovery, target validation, high-throughput screening, combinatorial chemistry, pharmaceutical productivity

Ashley Publications Ltd
www.ashley-pub.com

Introduction

Variations of the term 'chemical genomics' have been used to describe a diversity of approaches that address issues in target validation and drug discovery. The application most relevant to drug discovery encompasses an integrated approach, the objective of which is the rapid generation of drug-like 'tool' compounds that can be used to validate target-based therapeutic hypotheses. In practical terms, chemical genomics approaches often focus on a protein target class that can facilitate a parallel approach to multiple therapeutic indications and exploit synergies occurring throughout the drug discovery process. Chemical genomics approaches to drug discovery are evolving to deal with the unfavorable economics associated with coupling the high attrition rates encountered in early stages of drug discovery with the difficulty and expense of proving therapeutic target validity for biologically uncharacterized targets.

Postgenomic drug discovery

Modern drug discovery focuses on finding small organic chemicals that act *in vivo* at defined protein targets. Targets are selected through a variety of approaches that attempt to correlate a modulation of specific biochemical activity with a positive therapeutic effect. Once a target is selected from the estimated few thousand 'druggable' proteins in the human genome [1], an assay system is developed to allow screening [2,3] of chemical libraries to find active hits. Active hits are retested and frequently resynthesized to confirm activity. Confirmed hits are then modified through iterative rounds of chemical synthesis to produce a multiplicity of chemical analogs. If the analogs of the initial hits have improved potency, are judged by medicinal chemists to possess drug-like properties (related ultimately to *in vivo* bioavailability, duration of action, and lack of toxicity), and are related through a structure–activity relationship that indicates potential for additional property improvement, the hit is advanced to 'lead series' status. The lead series then undergoes further rounds of iterative chemical optimization, where improved compounds are subjected to an increasingly broad set of *in vitro* and *in vivo* tests in order to achieve desired levels of *in vivo* potency, desired pharmacokinetic properties, and minimization of adverse toxic effects. There is generally a high rate of attrition in the initial stages of drug discovery, with only a small fraction of initially screened targets resulting in compounds that advance to clinical development status.

The process of target selection implicitly assumes that a safe and specific modifier of the target protein will produce a therapeutically useful agent. However, as the pharmaceutical industry has moved from developing drugs against relatively well-understood protein targets, to a plethora of newer 'genomics' targets, the development of therapeutic proof of concept has become an increasingly important, if not dominant, risk factor in drug discovery. The pharmaceutical and biotechnology industries have developed a number of biology-based strategies

 ISSN 1462-2416

for target validation in order to increase the likelihood of therapeutic target relevance and improve overall attrition rates. Nevertheless, it is not yet clear whether the high cost of biology-based strategies applied early in discovery is repaid by improved economic efficiency in the overall drug discovery process. Chemical genomics approaches to drug discovery are evolving to deal with the unfavorable economics associated with coupling the high target attrition rates encountered in early stages of drug discovery with the difficulty and expense of proving therapeutic target relevance using indirect methods. Instead, chemical genomics approaches focus on efficient strategies for generating drug-like tool compounds that can be directly evaluated in biological models to test the therapeutic relevance of a target hypothesis.

The role of chemical genomics

The term 'chemical genomics' and related terms (e.g., chemical genetics, chemogenomics) have been applied to a diversity of approaches that use chemical compounds to probe biological systems [4,5]. While all of the approaches have at least peripheral relevance to drug discovery, the methods can be differentiated according to the extent to which they employ a stochastic versus directed approach to drug discovery, and how directly they are coupled to the development of a therapeutic proof of concept. The following discussion focuses on approaches that are most directly associated with the drug discovery process. We refer the reader to other reviews [4-7] for a more extensive discussion of the uses of chemical genomics tools for biological research.

Stochastic approaches

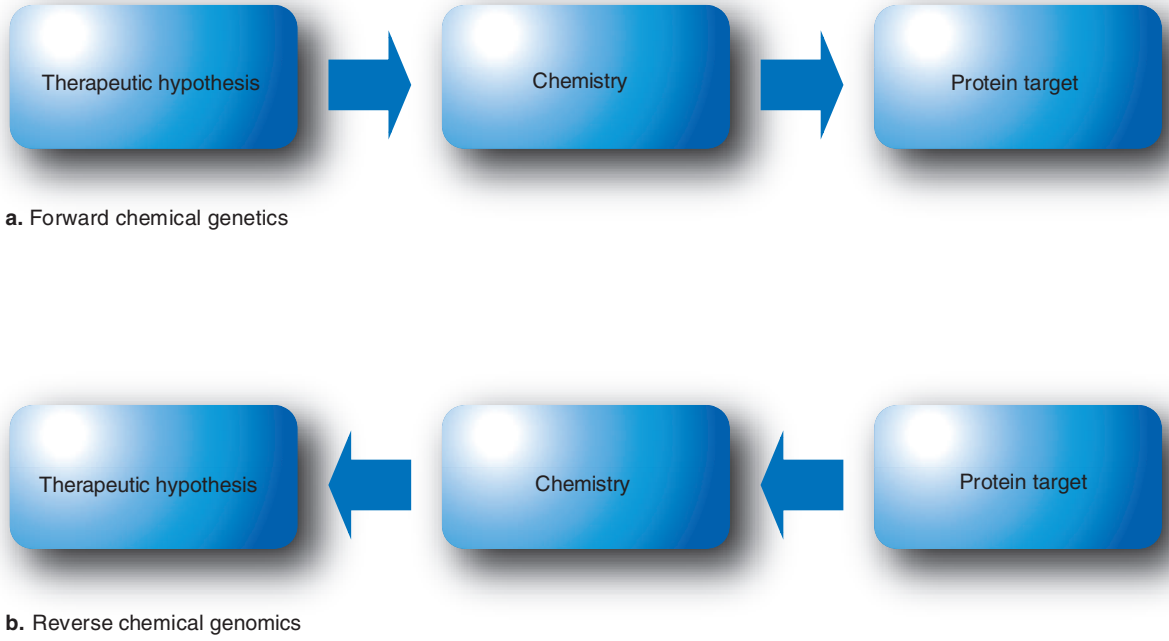
Stochastic approaches to biological discovery using chemistry are exemplified by 'chemical genetics' methods that examine the global biochemical, phenotypic, or genotypic response of a biological system when exposed to a foreign chemical agent [4-7]. In a fundamental sense, these methods are modern descendants of the classical drug discovery paradigm that screened natural products using *in vivo* biological models. This classical approach to drug discovery has produced many pharmaceutically useful agents in the absence of any knowledge of their specific molecular target. In the modern context of 'forward chemical genetics' [4] (**Figure 1**), the approach is enabled by powerful new tools that can identify specific molecular targets or correlate changes in a cellular phenotype or genotype to classify biochemical activity *in vivo* [8]. One manifestation of this approach contemplates the results of testing a large set of diverse chemical compounds against a correspondingly large subset of isolated proteins encoded by the human genome in order to find chemical compounds that can be used *in vivo* to investigate biochemical pathways [4-9]. Although a number of experimental approaches have been developed that allow high-throughput and direct measurements of drug binding to target proteins [3,10], there is a substantial component of serendipity attached to the collection of useful data using this approach. In addition, there are practical limitations on the expression and presentation of target proteins in high-throughput binding formats that can present potential ambiguities in interpretation of the results, particularly for proteins of unknown function or biochemical status (e.g., whether the protein target is activated through post-translational modification, the proper cofactors or allosteric effectors are present, binding properties are altered by protein or ligand immobilization chemistry etc.).

A second stochastic approach examines changes in gene expression that occur when normal, diseased, or engineered cell types are exposed to chemicals. Practical applications of this approach include the development of gene expression profiles that can be used to predict drug toxicology (due, for example to the upregulation of various liver enzymes) and the detection of other drug-related 'off-target' effects. There is optimism that by developing cellular expression fingerprints associated with compounds having known toxicological properties or mechanisms, it will be possible to predict the *in vivo* toxicological profiles of new compounds in advance of extensive animal toxicological testing [11,12] and to potentially develop higher throughput or predictive methods for general compound ADMET (absorption, distribution, metabolism, excretion, toxicology) properties. Although stochastic approaches using binding screens and or expression profiles as readouts can assist in target discovery, successful interpretation of the results requires substantial contextual information in order to assign a biochemically plausible role for the putative target, particularly if the results suggest a previously unknown biochemical function.

Target discovery using focused chemical approaches

More focused chemistry-based approaches to target discovery and/or validation probe tissue

Figure 1. Information flow in drug discovery.



a. Drug discovery strategies were historically motivated by a therapeutic hypothesis that could be screened in an intact biological disease model. Later advances in technology allowed definition of the ultimate protein drug target (a 'forward' chemical genetics paradigm). **b.** Genome sequencing work identifies protein targets that, when acted on by chemicals, may lead to a positive therapeutic outcome. Proving that an effector will produce a desired therapeutic effect by indirect means is a difficult process. The 'reverse' chemical genomics approach ultimately depends on finding a potent drug-like effector of the target that can validate the therapeutic hypothesis in a disease model.

level, localization or pathway organization associated with a specific class of potential drug target. Generally, the chemical probes are directed to a target class by virtue of specific reactive chemistry at the active or cofactor binding site. Several examples of this 'reverse' approach (**Figure 1**) have appeared that probe the expression level or cellular localization of enzymes generated during a pathological response as well as the role played by specific kinases in cellular signal transduction. Greenbaum *et al.* [13,14] report the use of fluorescently labeled epoxide inhibitor libraries in order to profile the level of cysteine proteases in complex cell extracts. Linking the inhibitor to biotin enabled the subsequent isolation of labeled proteins and their identification using mass spectrometry. Jassani *et al.* and Kidd *et al.* [15,16] pursue a similar objective by using biotinylated fluorophosphonate inhibitors to examine serine protease activity profiles in normal and invasive tumor cell lines. The former study identified well known tumor-associated proteases such as urokinase, as well as a novel membrane associated hydrolase as an important marker for tumor progression.

While the former approaches use chemical probes with broad activity against an enzyme class to provide evidence of expression level or tissue localization, approaches have also been developed to probe more specific biochemical functions, often by focusing on an individual member within a broad target class. Chan *et al.* [17] report the use of rapamycin, an immunosuppressive antibiotic, to study the global interactions and control functions of TOR, a protein kinase essential for cell growth. The strategy involved the measurement of the rapamycin sensitivity of over 2000 *Saccharomyces* strains for which individual proteins in the yeast genome were systematically deleted. The authors observed that over 100 mutants, clustering around eight broad cellular functions, exhibited the rapamycin sensitivity phenotype, indicating a wide range of roles for the TOR protein in cell regulation. Shokat and co-workers [18] have developed a hybrid approach to investigate the

specific roles of protein kinases, a family of enzymes with several hundred members [1] that play key regulatory roles in many cell signaling pathways. In this approach, the ATP binding site of the target kinase is modified through site-specific mutagenesis, so that it will function only in the presence of a cognately modified cofactor ATP analog. The approach achieves its generality through the relative structural similarity of the ATP binding site among kinases, which allows the introduction of the site-specific modification into different kinases as desired to study their function *in vivo*. The method has important applications in pathway-based target identification, and also in chemical genomics profiling to determine the effects of specific kinase inhibition on gene expression profiles.

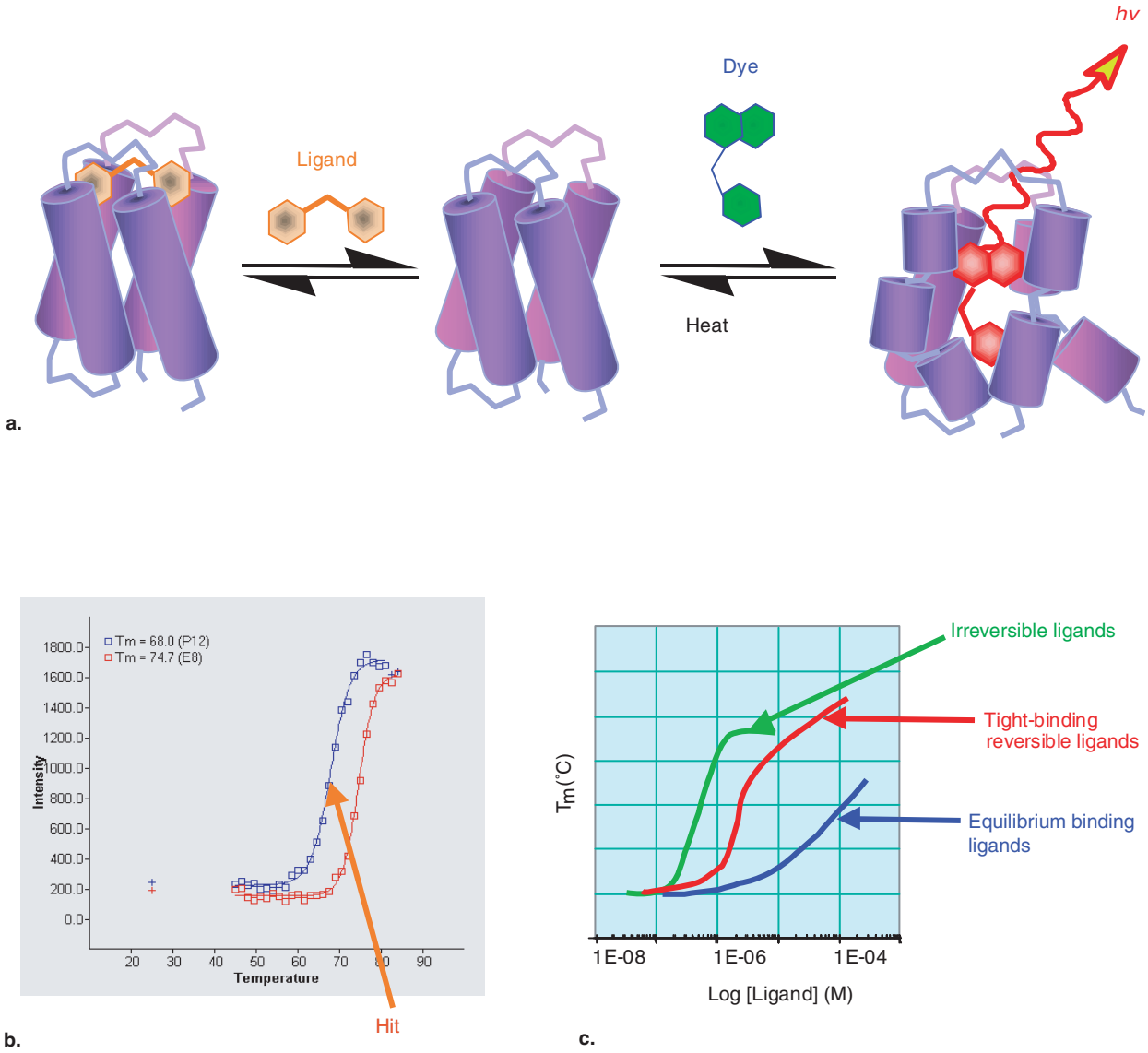
Chemical genomics as an integrated approach to drug discovery

The third application of chemical genomics focuses on the operational objective of rapidly discovering and optimizing chemical compounds to verify the 'druggability' and therapeutic validity of a drug target through direct testing in a suitable cellular or animal disease model. Unlike the approaches described above, which focus on aspects of target discovery and usually rely on a single experimental method, this broader definition of chemical genomics reflects an integrated approach that typically couples multiple technologies to rapidly identify and optimize drug-like compounds that can be used to test a therapeutic hypothesis *in vivo*. Although incorporating multiple technologies, chemical genomics approaches usually focus on a specific target class in order to exploit technical synergies and provide a parallel approach to the development and differentiation of drugs directed at multiple targets within the class [19].

There are three central components of an integrated chemical genomics platform. The first component is a molecular screen used to find initial hits, and ideally, also suitable for subsequent optimization of target binding affinity. Although classical drug screening programs often rely on enzymological, ligand displacement, or functional cellular assays to detect active compounds, many targets identified through genomics approaches are relatively uncharacterized biochemically. Quantitative ligand affinity methods are particularly suited to a chemical genomics approach for such targets because they do not rely on knowledge of the target biochemistry either to implement the screen or to subsequently optimize ligand-binding affinity. Many of the targets emerging from genomics studies are uncharacterized enzymes or members of receptor classes where it is unclear whether the target is able to bind small drug-like molecules with high affinity (i.e., whether they are 'druggable'). Direct binding assays provide a general way to screen for high affinity, small molecule drug leads, thus providing a direct test of 'druggability' for such targets. Direct affinity screens are generally based on the detection of some difference in the physical properties of a protein target when a screening compound is bound. There are several examples of high-throughput methods that have been used with some effectiveness in drug discovery programs. One method uses capillary electrophoresis to detect changes in the mobility of target protein in an applied electric field when it binds a ligand [101]. An alternative approach incubates the target protein with a mixture of compounds and uses a chromatographic column to separate the target protein together with any bound ligands. The target–ligand complexes are then heated to drive off the small molecules so that they can be analyzed and identified using mass spectroscopy [20,21]. This approach is particularly effective for screening hits from pooled compound libraries. A third approach measures the thermal stabilization that occurs when ligands bind to proteins [22]. The latter approach, which corresponds essentially to a miniaturized form of differential scanning calorimetry [23], requires somewhat more protein to conduct a screening campaign of several hundred thousand individual compounds than the capillary electrophoresis or mass spectroscopy methods. Nevertheless, the thermal shift method (**Figure 2**) has several unique advantages relative to other affinity screens. These include its very wide dynamic range, enabling its use for both initial screening and later stage lead optimization, and its utility as a secondary screen able to identify the molecular mechanism of inhibition [24].

A second required component of a modern chemical genomics platform is an efficient means of generating initial chemical screening libraries and subsequently optimizing hits obtained into drug-like compounds. Combinatorial chemistry, particularly as implemented through the parallel synthesis of individual compounds obtained by reacting diverse reagents with a common scaffold, has become a mainstay technology throughout pharmaceutical discovery. Nevertheless, to be effective, combinatorial

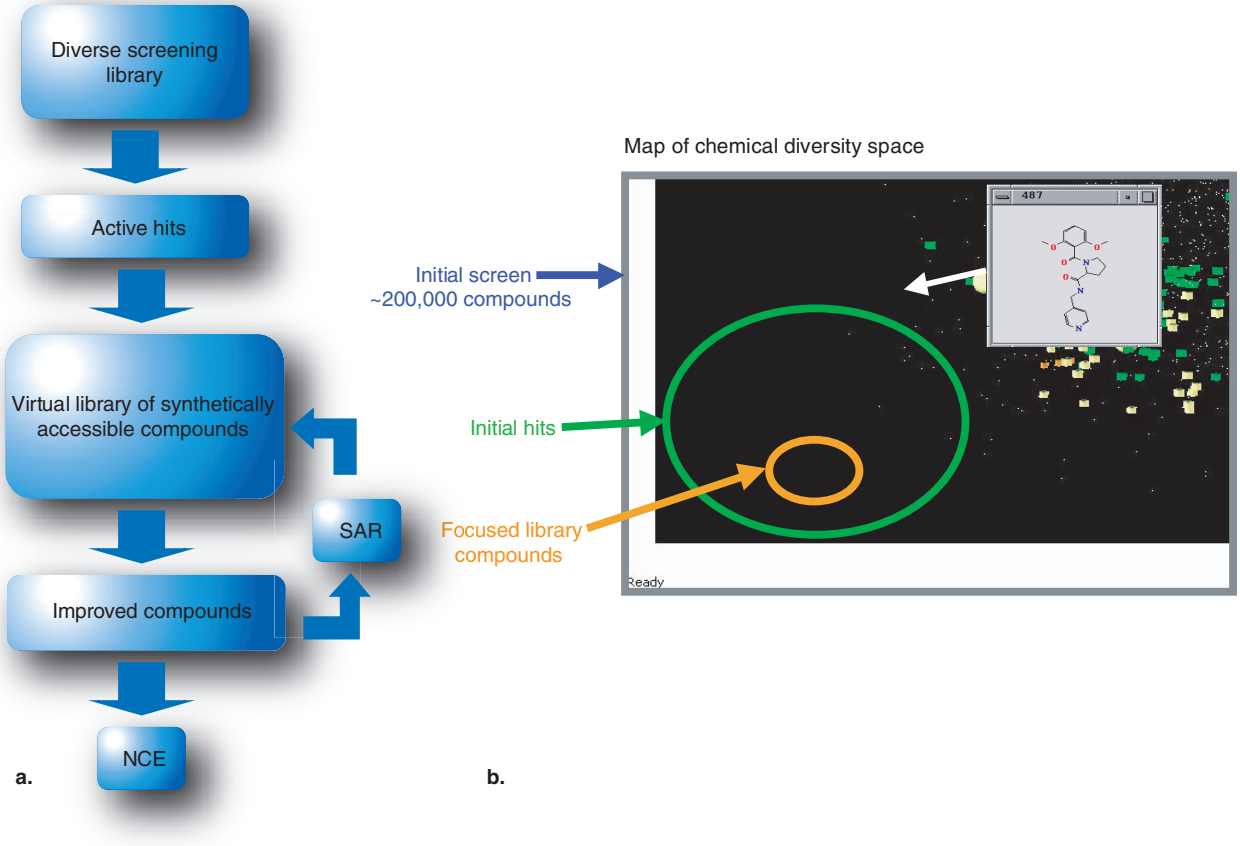
Figure 2. The thermal shift method.



a. Proteins unfold with a characteristic melting temperature T_m but are stabilized by the binding of ligands. The extent of stabilization or positive shift in T_m, is quantitatively related to the ligand binding constant over a dynamic range of K_a from picomolar to ~ 100 micromolar, and can be detected by a fluorescent probe dye that binds to the melted state of the protein and emits light (*hv*). **b.** Data output showing the appearance of a screening hit as a shifted T_m melting curve. **c.** Observations of the T_m as a function of ligand concentration can reveal details of the ligand binding mechanism, an important factor in advancing hits for subsequent chemical optimization.

chemistry has to be highly focused to find the few useful compounds among the vast number that are potentially accessible through random combinatorial synthesis (**Figure 3**). Given the scale of the accessible chemical space, there is a trade-off between generating libraries with the maximum drug-like diversity [25-28] versus libraries that are more focused on specific druggable target classes [29-32]. In either case, chemi-informatic computational approaches are required to tailor the design of combinatorial libraries so that they simultaneously co-optimize disparate properties required both for target potency and ultimate pharmaceutical development. Sophisticated chemi-informatic approaches can co-optimize diverse properties such as compound molecular

Figure 3. The figure illustrates how successive rounds of optimization were directed to converge in chemical property space.



a. Schema for hit discovery and lead optimization using combinatorial chemistry. **b.** Effective compound optimization requires a means of representing the relationships between chemical structures in multi-parameter spaces.
NCE: New chemical entity; SAR: Structure–activity relationship.

weight, computed solubility, density of hydrogen bond donor and acceptor groups, requirements to fit a predefined set of structural criteria (determined by either the target active site structure or pre-existing structure–activity data on active compounds), and heuristic models of compounds ADMET properties, as well as practical constraints on defining the most efficient matrix synthesis strategies [33]. In addition, an extensive set of peripheral chemi-informatics computational tools are required to monitor compound production, quality control, database registration, biological testing, and data mining activities that drive the iterative process of drug property optimization [33].

Structural biology [19,34] constitutes the final key component of an integrated chemical genomics platform. Recent years have seen the advent of numerous technological advances impacting virtually all aspects of the 3D structure determination of proteins using both X-ray crystallography and NMR spectroscopy. 3D structural data has been used both for *ab initio* drug design and to aid virtual screening of compound libraries using computational docking approaches [35]. Both crystallography and NMR have also been used to physically screen compounds through X-ray analysis of target crystal structures which have been infused with screening compounds [36], or by analysis of target NMR data that can detect compound binding in solution [37]. Structure-based design uses the 3D structure of the target protein as a basis for designing drug molecules *ab initio* [35], or more

usually, structural analysis of its complexes with drug leads as the basis for iterative chemical modification of compounds with improved properties [38]. X-ray crystallography has proved to be particularly useful in the design and optimization of combinatorial library strategies [39,40] for several of the major classes of druggable targets including proteases [41] and kinases [42]. The 3D structure of one family member, together with homology models of other family members, can furnish a template for library design applicable to the broader target class (**Figure 4**). As pharmaceutical companies increasingly focus on discovering drugs for newly identified members of established target families, there is increased emphasis on strategies that can enhance drug selectivity for a defined target versus related members of the family. Structural analysis and 3D homology modeling of multiple family members within a target family play an increasingly important role in defining synthetic strategies to produce drugs acting specifically at a defined target.

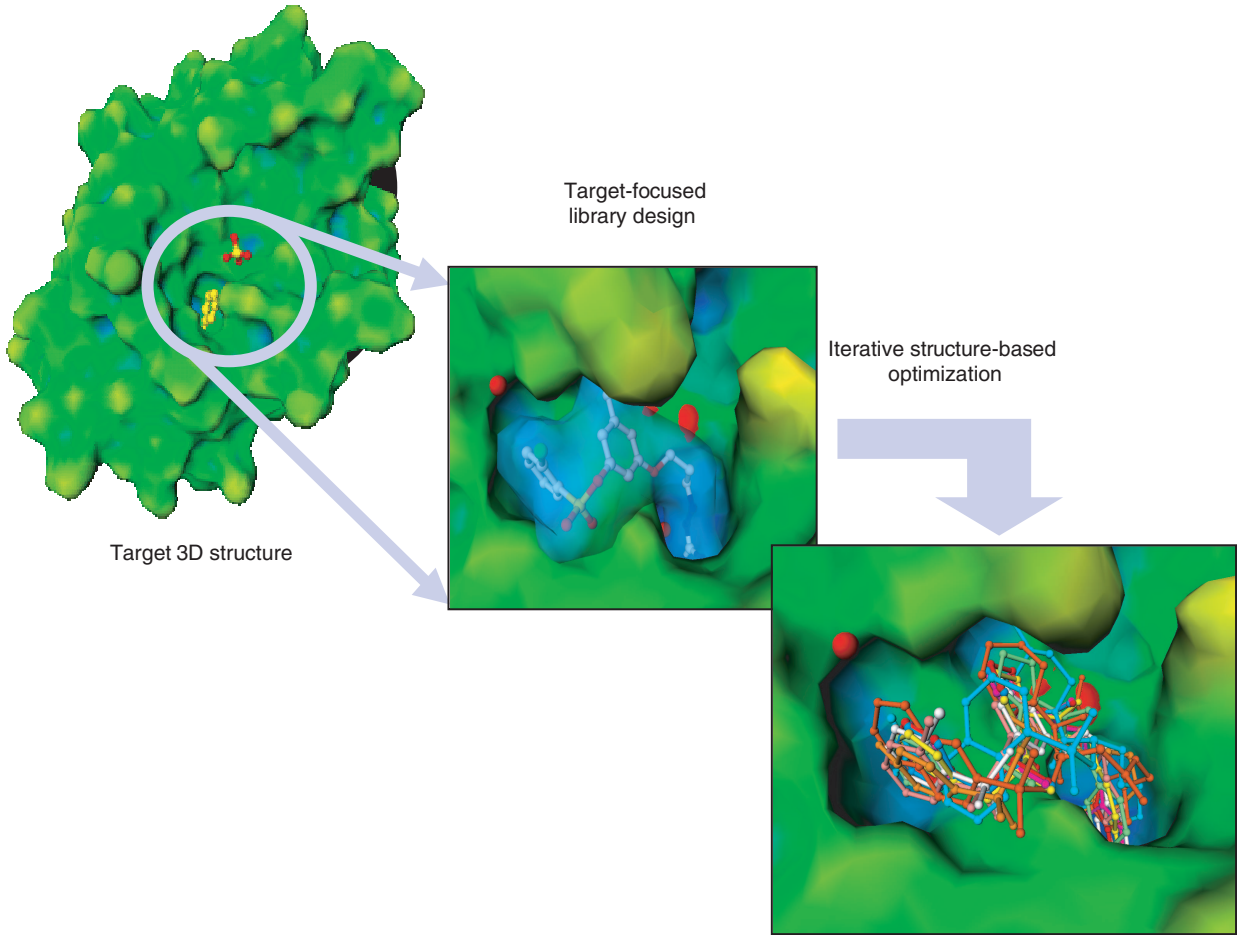
Expert opinion

Until recently, the pharmaceutical industry concentrated it efforts on ~ 500 protein targets with relatively well known biochemical functions and properties [43]. The revolution in genome sequencing that has occurred over the last decade has now defined virtually all of the potential targets for small molecule therapeutics [1]. This knowledge was initially envisioned to provide unprecedented opportunities for the discovery of new drugs with novel modes of therapeutic action. Nevertheless, this potential has yet to be realized and many reports now recognize that the overall productivity in the pharma industry has remained essentially constant over the last 10 years, despite a major increment in research and development investment [44-46]. Although there are many factors that impact pharma productivity, it is clear that much of the lost productivity can be attributed to the basic failure of new technology to make the discovery process more efficient. Although there has been a vast increase in the quantity of information obtained through genomics approaches and other technological advances, the translation of this information into effective discovery strategies (as opposed to interesting science) has proved elusive. Basically, these causes relate to the nature of 'genomics' targets, where relatively little may be known about the biochemical mechanism or biological function of a target protein, owing to its initial identification through upregulation in a disease context or from biological effects in a shotgun gene knockout experiment. As a consequence of this 'reverse' approach to drug discovery (**Figure 1**), a major preoccupation of the biotechnology and pharma industry has been to establish strategies by which targets can be biologically validated to demonstrate genuine therapeutic utility. Whatever the specific methodology, the biological validation process for novel targets is complex and expensive. Consequently, it is not surprising that the broad application of such methods in the early, high-risk stages of drug discovery, where the majority of target attrition occurs, leads to greatly increased discovery costs.

While the loss of effectiveness owing to the expenditure of resources in early target validation is difficult to quantify at the technical level, several publications [47-50] have recently appeared that test other basic assumptions that have motivated recent massive investments by pharma in drug discovery technology. For example, the trend towards very high-throughput screening with large chemical libraries seems to have fallen well short of expected productivity gains. The technical compromises (which frequently lead to screens being conducted under biochemically irrelevant conditions) and loss of precision that occur on adaptation of many assays for high-throughput platforms create high error rates that make hit follow-up by synthetic chemistry teams a very inefficient and frustrating process. The situation is exacerbated by the influx of targets from genomics where there is generally less known about basic target biochemistry, making the interpretation and relevance of screening results substantially more ambiguous than screens conducted on target classes whose biochemistry is better understood. There seems to be a major need to develop high quality secondary screening methods to sort out the genuinely useful hits obtained from large-scale screening campaigns and also to perform rapid biochemical characterization defining hit mechanism of action [24].

A final factor frustrating productivity concerns the need for better ways of managing and mining the vast quantity of very noisy data emerging from genomic and other high-throughput technologies [49]. The solution to this problem is complex owing to the diversity of the data encountered in the discovery process and the many difficulties associated with the integration of geographically distributed and disparate databases.

Figure 4. 3D structure determination provides a target class template for combinatorial library generation.



Subsequent 3D structure determination of drugs bound to the target can improve properties through recombination of features from different libraries or through conventional structure-based ligand design.

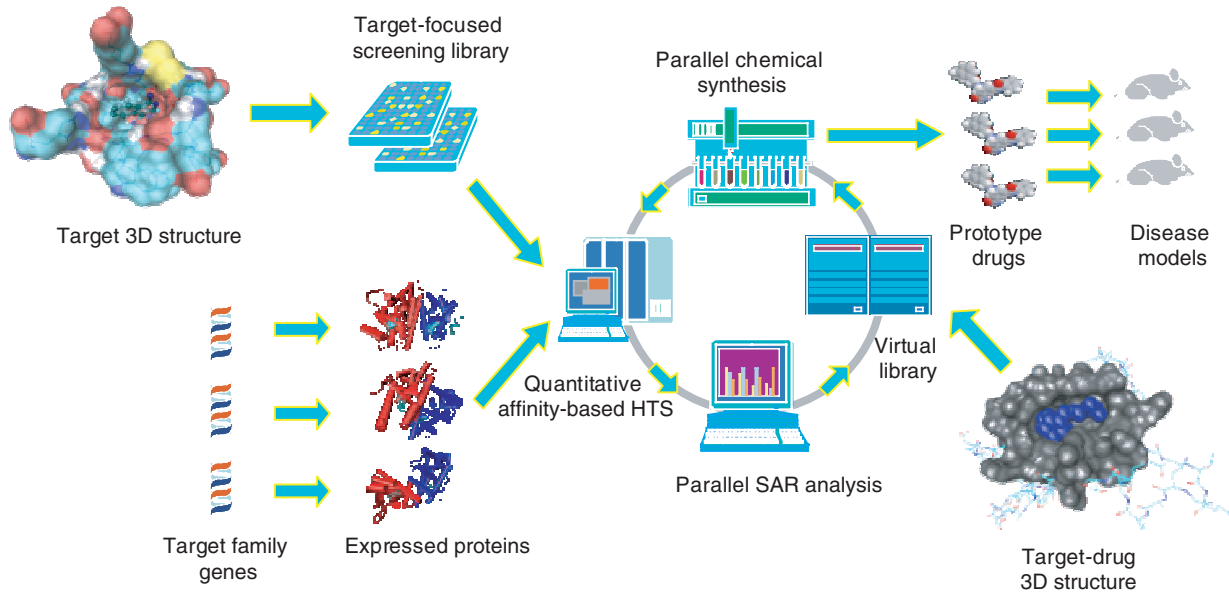
Outlook

Chemical genomics is evolving to address key issues posed by the sharp decrease pharmaceutical productivity that has occurred in the postgenomic era. This situation has arisen from the cost and difficulty of verifying the therapeutic utility of novel targets at the earliest stages of drug discovery where target attrition rates are high. An integrated chemical genomics approach focuses on the rapid generation of a potent drug-like compound that provides therapeutic target validation in a cellular or *in vivo* model of disease. This approach offers improvements in efficiency versus many current strategies for drug discovery, since it applies key biological resources at a stage in the process where many of the common attrition factors (e.g., target druggability, target specificity, cell permeability etc.) will have already been addressed.

An emerging paradigm for drug discovery in the postgenomic era

Figure 5 outlines a practical embodiment of a chemical genomics platform for drug discovery. The platform reflects a comprehensive approach to the discovery and optimization of multiple compounds interacting with different members of a druggable target family where the possibility exists to develop new therapeutic modalities. Well known examples include proteases [1,13-16,29], kinases [1,19,31], and nuclear hormone receptors [1,51,52]. 3D structural data from a representative target family member is used to computationally select or design chemical screening libraries for physical screening. Multiple target

Figure 5. An emerging paradigm for chemical genomics addressing a 'druggable' target family.



3D structural data from a representative target family member is used to select or design chemical screening libraries for physical screening. Multiple target family members, representing alternative therapeutic approaches or important counter screens, are cloned and expressed. A quantitative screening assay is used, in concert with a computer-directed combinatorial chemistry platform, to optimize and differentiate multiple drug-like compounds directed at different members of the target family. Structure-based design can play an important role in optimizing compound potency and specificity. The outputs of the process are multiple drug-like compounds with specific target selectivity that can be tested in multiple disease models.
HTS: High-throughput screening; SAR: Structure–activity relationship.

family members that potentially represent alternative therapeutic approaches or important counter screens are cloned and expressed in quantities sufficient to drive quantitative screening and 3D structural studies. The parallel cloning and expression of related molecules makes this process substantially more efficient and reliable on a per target basis than approaches addressing totally diverse targets. A high quality compound library is screened, using a few hundred thousand drug-like compounds selected for a combination of structural diversity and target focus. Screening is performed with a quantitative binding assay, which ideally has sufficient dynamic range to allow both initial hit screening (involving detection of compounds with K_a = ~ 10 µM) and to drive subsequent rounds of chemical compound optimization (producing compounds with K_a ~ nM). A quantitative binding assay providing direct thermodynamic readout [22-24] is able to find drugs that bind to target active or allosteric sites, to find drugs that bind cooperatively with small molecule, DNA, or protein substrates and cofactors, as well as to provide direct validation of molecular mechanisms of inhibition. Initial hits are optimized through chemical elaboration using computer-directed combinatorial strategies that are able to co-optimize target binding potency, specificity, and a host of other drug-like properties ultimately important for successful development. An efficient way to generate optimized compounds is to select them from a computer database of analogs of the initial screening library whose synthetic routes have been preverified and stored in the database. This allows the selected compounds to be synthesized and checked for quality using automated parallel synthesis and analysis methods. Depending on the stage of compound optimization, a variety of additional experimental tests can be made on the synthesized compounds and introduced into the computational optimization algorithm. These can include compound solubility, cell permeability, liver enzyme metabolic activity, toxicological expression profile data etc. 3D structural data for compounds bound to various targets can also be carried out in parallel, and this information used to design new libraries, improve target affinity and specificity, or other ADMET properties ultimately

Highlights

- Genome sequencing efforts have defined a druggable genome of a few thousand protein targets, predominantly clustered in biochemically or functionally related protein families.
- Drug discovery programs directed toward newer genomics targets present unusual difficulties owing both to a lack of basic biochemical knowledge and the difficulty of demonstrating therapeutic validation.
- Stochastic chemical genomics approaches probe the global response of a biological system on exposure to chemical compounds.
- Focused chemical genomics approaches use chemicals as detailed probes of biochemical pathways that can play a key role in target identification and validation.
- An integrated chemical genomics platform uses affinity-based screening, directed combinatorial chemistry, and structure-based drug design to rapidly develop drug-like tool compounds that can validate a target-based therapeutic hypothesis *in vivo*.

important for successful drug development [53]. The outputs of the process are multiple drug-like compounds with specific target selectivity that can be tested in multiple disease models.

The integrated chemical genomics process keys from the identification of a DNA sequence inferred to have therapeutic significance in a disease context. The basic information furnished by the target sequence facilitates target protein synthesis and the subsequent discovery and optimization of drug-like 'tool compounds' that interact with the target and can be evaluated in a biological disease model to directly test a therapeutic hypothesis. The approach can be implemented as a highly parallel process, and is particularly well suited to the discovery of drugs in broad families, where inter-target specificity may be a critical factor in the ultimate development of therapeutic agents with minimum side effects. The chemical genomics approach defers investment in biological target validation to a later stage in the discovery cycle, where resources can be deployed more efficiently and with a higher probability of success, thus providing a more direct route to efficiently finding new drugs.

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Hopkins AL, Groom CR: The druggable genome. *Nature Rev. Drug Disc.* 1, 727-730 (2002).

• **A realistic, albeit conservative, view of the number of new drug targets emerging from the human genome sequencing effort.**

2. Houston JG, Banks M: The chemical-biological interface: developments in automated and miniaturized screening technology. *Curr. Opin. Biotechnol*. 8(6), 734-740 (1997).
3. Croston GE: Functional cell-based uHTS in chemical genomic drug discovery. *Trends Biotechnol.* 20(3), 110-115 (2002).
4. Schreiber SL: The small molecule approach to biology. *Chem. Eng. News* 81(9), 51-61 (2003).

• **A recent review of work principally from the Schreiber lab giving the academic perspective on the relationship between chemistry and biology.**

5. Schreiber SL: Chemical genetics resulting from a passion for synthetic organic chemistry. *Bioorg. Med. Chem*. 6, 1127-1152 (1998)
6. Zheng SXF, Chan TF: Chemical genomics in the global study of protein functions. *DDT* 7(3), 197-205 (2002).
7. Zheng SXF, Chan TF: Chemical genomics: a systematic approach in biological research and drug discovery. *Curr. Issues. Mol. Biol.* 4, 33-43 (2002).
8. Watterson DM, Haiech J, Van Eldik LJ: Discovery of new chemical classes of synthetic ligands that suppress neuroinflammatory responses. *J. Mol. Neurosci*. 19, 89-93 (2002).
9. Alaimo PJ, Shogren-Knaak MA, Shokat KM *et al.*: Chemical genetic approaches for the elucidation of signaling pathways. *Curr. Opin. Chem. Biol.* 5, 360-367 (2001).
10. Zimmermann B, Hahnefeld C, Herberg F W: Applications of biomolecular interaction analysis in drug development. *Targets* 1(2), 66-73 (2002).
11. Browne LJ, Furness LM, Natsoulis G, Pearson C, Jarnagin K: Chemiogenomics – pharmacology with genomics tools. *Targets* 1(2), 59-65 (2002).
12. Staunton JE, Slonim DK, Coller HA *et al.*: Chemosensitivity prediction by transcriptional profiling. *Proc. Natl. Acad. Sci. USA* 98(19), 10787-10792 (2001).
13. Greenbaum D, Medzihradszky KF, Burlingame A, Bogyo M: Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol.* 7, 569-581 (2000).
14. Greenbaum D, Brauch A, Hayrapetian L *et al.*: Chemical Approaches for Functionally Probing the Proteome. *Mol. Cell. Proteomics* 1(1), 60-68 (2002).
15. Jessani N, Liu Y, Humphrey M, Cravatt BF: Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. *Proc. Natl. Acad. Sci. USA* 99(16), 10335-10340 (2002).
16. Kidd, Liu Y, Cravatt BF: Profiling serine hydrolase activities in complex proteomes. *Biochemistry* 40, 4005-4015 (2001).
17. Chan TF, Carvalho J, Riles L, Zheng SXF: A chemical genomics approach toward understanding the global functions of the target of rapamycin protein (TOR). *Proc. Natl. Acad. Sci. USA* 97(24), 13227-13232 (2000).
18. Shokat K, Velleca M: Novel chemical genetic approaches to the discovery of signal transduction inhibitors. *Drug Discov. Today* 7(16), 872-879 (2002).

•• **An interesting integration of chemistry and structure-directed protein engineering to enable a unique approach to in vivo characterization of kinase function.**

19. Caron P, Mullican MD, Mashal RD *et al.*: Chemogenomic approaches to drug discovery. *Curr. Opin. Chem. Biol.* 5, 464-470 (2001).
20. Lenz GR, Nash H M, Jindal S: Chemical ligands, genomics and drug discovery. *Drug Discov. Today* 5(4), 145-156 (2000).
21. Alaoui-Ismaili MH, Lomedico PT, Jindal S: Chemical genomics: discovery of disease genes and drugs. *Drug Discov. Today* 7(5), 292-294 (2002).
22. Pantoliano MP, Petrella E, Kwasnoski J *et al.*: High density miniaturized thermal shift assay as a general strategy for drug

discovery. *J. Biomol. Screening* 6(6), 492-440 (2001).

23. Plotnikov V, Rochalski A, Brandts M *et al.*: An autosampling differential scanning calorimeter instrument for studying molecular interactions. *ASSAY Drug Development Technologies* 1(1-1), 83-90 (2002).
24. Todd M, Salemme FR: Direct binding assays for pharma screening. *Genetic Engineering News* 23(3), 28-29 (2003).

•• **A discussion of the use of high-throughput thermodynamic methods to elucidate biochemical mechanisms of hits obtained from compound screening with novel genomics targets.**

25. Weber L: High-diversity combinatorial libraries. *Curr. Opin. Chem. Biol.* 4, 295-302 (2000).
26. Veber DF, Drake FH, Gowen M: The new partnership of genomics and chemistry for accelerated drug development. *Curr. Opin. Chem. Biol.* 1, 151-156 (1997).
27. Dorman G, Krajcs P, Darvas F: Chemical library approaches to target validation in the post-genomics era. *Curr. Drug Disc*. 21-24 (2001).
28. Namchuk M: Finding the molecules to fuel chemogenomics. *Targets* 1(4), 125-127 (2002).
29. Dhanoa DS, Soll RM, Subasinghe N *et al.*: Serine proteases-directed small molecule probe libraries. *Med. Chem. Res.* 8(4/5), 187-205 (1998).
30. Auld DS, Diller D, Ho KK: Targeting signal transduction with large combinatorial collections. *Drug Discov. Today* 7(24), 1206-1213 (2002).
31. Gray NS, Wodicka L, Thunnisssen AWH *et al.*: Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* 281, 533-538 (1998).
32. Sawyers C: Finding the next Gleevec: FLT3 targeted kinase inhibitor therapy for acute myeloid leukemia. *Cancer Cell* 1, 413-415 (2002).
33. Agrafiotis D, Lobannov V, Salemme FR: Combinatorial informatics in the post-genomics era. *Nature Drug Discovery* 1, 338-346 (2002).

• **A comprehensive discussion of the application of chemi-informatics for iterative multi-objective optimization of drug properties with combinatorial chemistry.**

34. Murko M, Caron P: Transforming the genome to drug discovery. *Drug Discov. Today* 7, 583-584 (2002).

• **A pioneering treatment of the role of structural biology and computational modeling for drug discovery in large target classes.**

35. Zanders ED, Bailey DS, Dean PM: Probes for chemical genomics by design. *Drug Discov. Today* 7(13), 711-718 (2002).
36. Carr R, Jhoti H: Structure-based screening of low-affinity compounds. *Drug Discov. Today* 7(9), 522-527 (2002).
37. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW: SAR by NMR: a method for discovering high affinity ligands for proteins. *Science* 274, 1531-1534 (1996).
38. Bailey D, Brown D: High-throughput chemistry and structure-based design: survival of the smartest. *Drug Discov. Today* 6, 57-59 (2001).
39. Salemme FR, Spurlino J, Bone R: Serendipity meets precision: the integration of structure-based drug design and combinatorial chemistry for efficient drug discovery. *Structure* 5, 319-324 (1997).

• **An early discussion addressing the synergy between combinatorial chemistry and structure-based drug design.**

40. Bone R, Salemme FR: The integration of structure-based design and directed combinatorial chemistry for new pharmaceutical discovery. In: *Structure-Based Drug Design.* Veerapandian P (Ed.), Marcel Dekker, Inc., New York, 525-539 (1997).
41. Illig C, Eisennagel S, Bone R *et al.*: Expanding the envelope of structure-based drug design using chemical libraries: application to small-molecule inhibitors of thrombin. *Med. Chem. Res.* 8(4/5) 244-260 (1998).
42. Scapin G: Structural biology in drug design: selective protein kinase inhibitors. *Drug Discov. Today* 7(11), 601-611 (2002).
43. Drews J: A historical perspective. *Science* 287, 1960-1964 (2000).
44. Branca M: The trouble with pharmaceutical innovation. *Bio-IT World* 32-34 (2002).
45. Canada SK, Caldwell C, Caldwell JS: Fulfilling the promise: drug discovery in the post-genomic era. *Drug Discov. Today* 8(4), 168-174 (2003).
46. Frantz S, Smith A: New drug approvals for 2002. *Nat. Rev. Drug Disc.* 2, 95-96 (2003).
47. Fox S, Wang H, Sopchak L, Farr-Jones S: High throughput screening 2002: moving toward increased success rates. *J. Biomol. Scr.* 7(4), 313-316 (2002).
48. Sills MA, Weiss D, Pharm Q, Schweitzer R, Wu X, Wu JJ: Comparison of assay technologies for a tyrosine kinase assay generates different results in high throughput screening. *J. Biomol. Scr.* 7(3), 191-214 (2002).

•• **One of the few objective analyses illustrating the error rates and inconsistencies in hit results obtained from 'industry-standard' high-throughput screens.**

49. Fischer HP: Turning quantity in quality: novel quality assurance strategies for data produced by high-throughput genomics technologies. *Targets* 1(4), 139-146 (2002)
50. Chrétien JR: Efficiency control in drug discovery. *Curr. Drug Disc.* 109-112 (2001)
51. Willson TM, Moore JT: Genomics versus orphan nuclear receptors – a half-time report. *Mol. Endocrinology* 16(6), 1135-1144 (2002).
52. Willson TM, Jones SA, Moore JT, Kliewer SA: Chemical genomics: functional analysis of orphan nuclear receptors in the regulation of bile acid metabolism. *Med. Res. Rev.* 21(6), 513-522 (2001).
53. Landro JA, Taylor ICA, Stirtan WG *et al.*: HTS in the new millennium The role of pharmacology and flexibility. *J. Pharmacol Toxicol. Methods* 44, 273-289 (2000).

Website

101. www.cetek.com