A REVIEW ON COMBINATORIAL CHEMISTRY IN DRUG DISCOVERY

A. Elumalai* and M. Chinna Eswariah

*Anurag Pharmacy College, Ananthagiri (v), Kodad (M), Nalgonda (Dt), Andhra Pradesh, India, 508-206.

ABSTRACT

Combinatorial chemistry involves the rapid synthesis or the computer simulation of a large number of different but often structurally related molecules or materials. In a combinatorial synthesis, the number of compounds made increases exponentially with the number of chemical steps. In a binary light-directed synthesis, 2n compounds can be made in n chemical steps. Combinatorial chemistry is especially common in CADD (Computer aided drug design) and can be done online with web based software, such as Molinspiration. The present review focused on synthesis of molecules in a combinatorial fashion can quickly lead to large numbers of molecules.

Keywords: Combinatorial chemistry, Lead molecule, Applications.

INTRODUCTION

Synthesis of molecules in a combinatorial fashion can quickly lead to large numbers of molecules. For example, a molecule with three points of diversity (R1, R2, and R3) can generate possible structures, where and are the numbers of different substituents utilized. Although combinatorial chemistry has only really been taken up by industry since the 1990s, its roots can be seen as far back as the 1960s when a researcher at Rockefeller University, Bruce Merrifield, started investigating the solid-phase synthesis of peptides. Professor Pieczenik, a colleague of Nobel Laureate Merrifield, synthesized the first combinatorial library. US Patent 5,866,363. In the 1980s researcher H. Mario Geysen developed this technique further, creating arrays of different peptides on separate supports, but not a combinatorial library based on random synthesis.

In its modern form, combinatorial chemistry has probably had its biggest impact in the pharmaceutical industry. Researchers attempting to optimize the activity profile of a compound create a 'library' of many different but related compounds. Advances in robotics have led to an industrial approach to combinatorial synthesis, enabling companies to routinely produce over 100,000 new and unique compounds per year. In order to handle the vast number of structural possibilities, researchers often create a 'virtual library', a computational enumeration of all possible structures of a given pharmacophore with all available reactants. Such a library can consist of thousands to millions of 'virtual' compounds. The researcher will select a subset of the 'virtual library' for actual synthesis, based upon various calculations and criteria [1,2].

Materials Science

Materials science has applied the techniques of combinatorial chemistry to the discovery of new materials. This work was pioneered by P.G. Schultz et al. in the mid nineties in the context of luminescent materials obtained by co-deposition of elements on a silicon substrate. His work was preceded by J. J. Hanak in 1970 but the computer and robotics tools were not available for the method to spread at the time. Work has been continued by several academic groups as well as companies with large research and development programs. The technique has been used extensively for catalysis, coatings, electronics, and many other fields. The application of appropriate informatics tools is critical to handle, administer, and store the vast volumes of data produced. New types of Design of experiments methods have also been developed to

Corresponding Author:- A. Elumalai Email:- malairx@gmail.com

efficiently address the large experimental spaces that can be tackled using combinatorial methods.

Diversity-oriented libraries

Even though combinatorial chemistry has been an essential part of early drug discovery for more than two decades, so far only one de novo combinatorial chemistry-synthesized chemical has been approved for clinical use by FDA (sorafenib, a multikinase inhibitor indicated for advanced renal cancer). The analysis of poor success rate of the approach has been suggested to connect with the rather limited chemical space covered by products of combinatorial chemistry. When comparing the properties of compounds in combinatorial chemistry libraries to those of approved drugs and natural products, Feher and Schmidt noted that combinatorial chemistry libraries suffer particularly from the lack of chirality, as well as structure rigidity, both of which are widely regarded as drug-like properties. Even though natural product drug discovery has not probably been the most fashionable trend in pharmaceutical industry in recent times, a large proportion of new chemical entities still are nature-derived compounds, and thus, it has been suggested that effectiveness of combinatorial chemistry could be improved by enhancing the chemical diversity of screening libraries. As chirality and rigidity are the two most important features distinguishing approved drugs and natural products from compounds in combinatorial chemistry libraries, these are the two issues emphasized in so-called diversity oriented libraries, i.e. compound collections that aim at coverage of the chemical space, instead of just huge numbers of compounds.

Dynamic combinatorial chemistry

Dynamic combinatorial chemistry (DCC) is a method to the generation of new molecules formed by reversible reaction of simple building blocks under thermodynamic control [1]. The library of these reversibly interconverting building blocks is called dynamic combinatorial library (DCL). All constituents in a DCL are in equilibrium, and their distribution is determined by their thermodynamic stability within the DCL. The interconversion of these building blocks may involve covalent or non-covalent interactions. When a DCL is exposed to an external influence (such as proteins or nucleic acids), the equilibrium shifts and those components that interact with the external influence are stabilised and amplified, allowing more of the active compound to be formed.

Protein-directed dynamic combinatorial chemistry

One of the key areas for development within the DCC field is the use of proteins to influence the evolution and generation of components within a DCL.[3] Protein-directed DCC provides a way to generate, identify and rank novel protein ligands, and therefore have huge potential in enzyme inhibitor and drug discovery. However the development of protein-directed DCC has not been straightforward, as the reversible reaction must occur in aqueous solution at biological pH and temperature, the components of the DCL must be compatible with proteins, and the DCC system must be amenable to efficient screening.

Reversible reactions for protein-directed DCC

Several reversible reactions have been applied in protein-directed DCC. These included boronate ester formation, imine formation, hydrazone formation, disulphide formation, thiol-enone exchange and hemithiolacetal formation.

Analytical methods for protein-directed DCC

Several analytical techniques have been applied to the analysis of protein-directed DCL. These include HPLC, mass spectrometry, NMR spectroscopy, and X-ray crystallography.

Other Applications

DCC is useful in identifying molecules with unusual binding properties, and provides synthetic routes to complex molecules that aren't easily accessible by other means. These include smart materials, foldamers, self-assembling molecules with interlocking architectures and new soft materials. Recently, the application of DCC to detect volatile bioactive compounds, i.e. the amplification and sensing of scent, was proposed in a concept paper.

In biotechnology, combinatorial biology is the creation of a large number of compounds (usually proteins or peptides) through technologies such as phage display. Similar to combinatorial chemistry, compounds are produced by biosynthesis rather than organic chemistry. This process was developed independently by Richard A. Houghten and H. Mario Geysen in the 1980s. Combinatorial biology allows the generation and selection of the large number of ligands for high-throughput screening.

These large numbers of peptides are generated and screened by physically linking a gene encoding a protein and a copy of this protein. This could involve the protein being fused to the M13 minor coat protein pIII, with the gene encoding this protein being held within the phage particle. Large libraries of phages with different proteins on their surfaces can then be screened through automated selection and amplification for a protein that binds tightly to a particular target.

High-throughput screening

High-throughput screening (HTS) is a method for scientific experimentation especially used in drug discovery and relevant to the fields of biology and chemistry. Using robotics, data processing and control software, liquid handling devices, and sensitive detectors,

High-Throughput Screening allows a researcher to quickly conduct millions of chemical, genetic or pharmacological tests. Through this process one can rapidly identify active compounds, antibodies or genes which modulate a particular biomolecular pathway. The results of these experiments provide starting points for drug design and for understanding the interaction or role of a particular biochemical process in biology.

Assay plate preparation

The key labware or testing vessel of HTS is the microtiter plate: a small container, usually disposable and made of plastic that features a grid of small, open divots called wells. Modern microplates for HTS generally have either 384, 1536, or 3456 wells. These r all multiples of 96, reflecting the original 96 well microplate with 8 x 12 9mm spaced wells. Most of the wells contain experimentally useful matter, depending on the nature of the experiment. This could be an aqueous solution of dimethyl sulfoxide (DMSO) and some other chemical compound, the latter of which is different for each well across the plate. It could also contain cells or enzymes of some type [3,4].

A screening facility typically holds a library of stock plates, whose contents are carefully catalogued, and each of which may have been created by the lab or obtained from a commercial source. These stock plates themselves are not directly used in experiments; instead, separate assay plates are created as needed. An assay plate is simply a copy of a stock plate, created by pipetteing a small amount of liquid (often measured in nanoliters) from the wells of a stock plate to the corresponding wells of a completely empty plate.

Reaction observation

o prepare for an assay, the researcher fills each well of the plate with some logical entity that he or she wishes to conduct the experiment upon, such as a protein, or an animal embryo. After some incubation time has passed to allow the biological matter to absorb, bind to, or otherwise react with the compounds in the wells, measurements are taken across all the plate's wells, either manually or by a machine. Manual measurements are often necessary when the researcher is using microscopy to seek changes or defects in embryonic development caused by the wells' compounds, looking for effects that a computer could not easily determine by itself. Otherwise, a specialized automated analysis machine can run a number of experiments on the wells (such as shining polarized light on them and measuring reflectivity, which can be an indication of protein binding). In this case, the machine outputs the result of each experiment as a grid of numeric values, with each number mapping to the value obtained from a single well. A high-capacity analysis machine can measure dozens of plates in the space of a few minutes like this, generating thousands of experimental datapoints very quickly.

Depending on the results of this first assay, the researcher can perform follow up assays within the same screen by cherrypicking liquid from the source wells that gave interesting results (known as hits) into new assay plates, and then re-running the experiment to collect further data on this narrowed set, confirming and refining observations.

Automation systems

Automation is an important element in HTS's usefulness. Typically, an integrated robot system consisting of one or more robots transports assay-microplates from station to station for sample and reagent addition, mixing, incubation, and finally readout or detection. An HTS system can usually prepare, incubate, and analyze many plates simultaneously, further speeding the data-collection process. HTS robots currently exist which can test up to 100,000 compounds per day. Automatic colony pickers pick thousands of microbial colonies for high throughput genetic screening. The term uHTS or ultra high throughput screening refers (circa 2008) to screening in excess of 100,000 compounds per day [4,5].

Experimental design and data analysis

With the ability of rapid screening of diverse compounds (such as small molecules or siRNAs) to identify active compounds, HTS has led to an explosion in the rate of data generated in recent years. Consequently, one of the most fundamental challenges in HTS experiments is to glean biochemical significance from mounds of data, which relies on the development and adoption of appropriate experimental designs and analytic methods for both quality control and hit selection. HTS research is one of the fields which have a feature described by Eisenstein as follows: soon, if a scientist does not understand some statistics or rudimentary data-handling technologies, he or she may not be considered to be a true molecular biologist and thus will simply become a dinosaur.

Quality control

High-quality HTS assays are critical in HTS experiments. The development of high-quality HTS assays requires the integration of both experimental and computational approaches for quality control (QC). Three important means of QC are (i) good plate design, (ii) the selection of effective positive and negative chemical/biological controls, and (iii) the development of effective QC metrics to measure the degree of differentiation so that assays with inferior data quality can be identified. A good plate design helps to identify systematic errors (especially those linked with well position) and determine what normalization should be used to remove/reduce the impact of systematic errors on both QC and hit selection.

Effective analytic QC methods serve as a gatekeeper for excellent quality assays. In a typical HTS experiment, a clear distinction between a positive control and a negative reference such as a negative control is an index for good quality. Many quality assessment measures have been proposed to measure the degree of differentiation between a positive control and a negative reference. Signal-to-background ratio, signal-to-noise ratio, signal window, assay variability ratio, and Z-factor have been adopted to evaluate data quality. Strictly standardized mean difference (SSMD) has recently been proposed for assessing data quality in HTS assays [5,6].

Hit selection

A compound with a desired size of effects in an HTS screen is called a hit. The process of selecting hits is called hit selection. The analytic methods for hit selection in screens without replicates (usually in primary screens) differ from those with replicates (usually in confirmatory screens). For example, the z-score method is suitable for screens without replicates whereas the t-statistic is suitable for screens with replicate. The calculation of SSMD for screens without replicates also differs from that for screens with replicates.

For hit selection in primary screens without replicates, the easily interpretable ones are average fold change, mean difference, percent inhibition, and percent activity. However, they do not capture data variability effectively. The z-score method or SSMD, which can capture data variability based on an assumption that every compound has the same variability as a negative reference in the screens. However, outliers are common in HTS experiments, and methods such as z-score are sensitive to outliers and can be problematic. Consequently, robust methods such as the z-score method, SSMD, B-score method, and quantile-based method have been proposed and adopted for hit selection.

In a screen with replicates, we can directly estimate variability for each compound; consequently, we should use SSMD or t-statistic that does not rely on the strong assumption that the z-score and z*-score rely on. One issue with the use of t-statistic and associated p-values is that they are affected by both sample size and effect size. They come from testing for no mean difference, thus are not designed to measure the size of compound effects. For hit selection, the major interest is the size of effect in a tested compound. SSMD directly assesses the size of effects. SSMD has also been shown to be better than other commonly used effect sizes. The population value of SSMD is comparable across experiments and thus we can use the same cutoff for the population value of SSMD to measure the size of compound effects.

Techniques for increased throughput and efficiency

Unique distributions of compounds across one or many plates can be employed to increase either the number of assays per plate, or to reduce the variance of assay results, or both. The simplifying assumption made in this approach is that any N compounds in the same well will not typically interact with each other, or the assay target, in a manner that fundamentally changes the ability of the assay to detect true hits.

For example, imagine a plate where compound A is in wells 1-2-3, compound B is in wells 2-3-4, and compound C is in wells 3-4-5. In an assay of this plate against a given target, a hit in wells 2, 3, and 4 would indicate that compound B is the most likely agent, while also providing three measurements of compound B's efficacy against the specified target. Commercial applications of this approach involve combinations in which no two compounds ever share more than one well, to reduce the (second-order) possibility of interference between pairs of compounds being screened.

Recent advances

Faster screening (100 million reactions in 10 hours) at 1 millionth the cost (using $10-7$ times the reagent volume) than conventional techniques using drop-based microfluidics. Drops of fluid separated by oil replace microplate wells and allow analysis and hit sorting while reagents are flowing through channels. In 2010 researchers developed a silicon sheet of lenses that can be placed over microfluidic arrays to allow the fluorescence measurement of 64 different output channels simultaneously with a single camera. This process can analyze 200,000 drops per second [6,7].

Increasing lab utilization of HTS

HTS is a relatively recent innovation, made lately feasible through modern advances in robotics and high-speed computer technology. It still takes a highly specialized and expensive screening lab to run an HTS operation, so in many cases a small-to-moderately sized research institution will use the services of an existing HTS facility rather than set up one for it.

There is a trend in academia to be their own drug discovery enterprise. (High-throughput screening goes to school) These facilities, which normally are only found in industry, are now increasingly be found as well at universities. UCLA for example, features an HTS laboratory (Molecular Screening Shared Resources (MSSR, UCLA) which can screen more than 100,000 compounds a day on a routine basis. The University of Illinois also has a facility for HTS, as does the University of Minnesota. The Rockefeller University, has an open access (infrastructure) HTS Resource Center HTSRC (The Rockefeller University, HTSRC) which offers a library of over 165,000 compounds. Northwestern University's High Throughput Analysis Laboratory supports target identification, validation, assay development, and compound screening.

In the United States, the National Institute of Health or NIH has created a nationwide consortium of small molecule screening centers that has been recently funded to produce innovative chemical tools for use in biological research. The Molecular Libraries Screening Center Network or MLSCN performs HTS on assays provided by the research community, against a large library of small molecules maintained in a central molecule repository. Early drug discovery involves several phases from target identification to preclinical development. The identification of small molecule modulators of protein function and the process of transforming these into high-content lead series are key activities in modern drug discovery.[1] The Hit-to-Lead phase is usually the follow-up of high-throughput screening (HTS) [7,8].

Hit confirmation

The Hit confirmation phase will be performed during several weeks as follows:
Re-testing: compounds that were found active against the selected target are re-tested using the same assay conditions used during the HTS. Dose response curve generation: several compound concentrations are tested using the same assay, an IC50 or EC50 value is then generated. Methods are being developed that may allow the reuse of the compound that generated the hit in the initial HTS step. These molecules are removed from beads and transferred to a microarray for quantitative assessment of binding affinities in a seamless approach that could allow for the investigation of more hits and larger libraries. Orthogonal testing: Confirmed hits are assayed using a different assay which is usually closer to the target physiological condition or using a different technology. Secondary screening: Confirmed hits are tested in a functional assay or in a cellular environment. Membrane permeability is usually a critical parameter. Chemical amenability: Medicinal chemists evaluate compounds according to their synthesis feasibility and other parameters such as up-scaling or costs. Intellectual property evaluation: Hit compound structures are quickly checked in specialized databases to define patentability.

Biophysical testing

Nuclear magnetic resonance (NMR), Isothermal Titration Calorimetry, dynamic light scattering, surface plasmon resonance, dual polarisation interferometry, microscale thermophoresis (MST) are commonly used to assess whether the compound binds effectively to the target, the stoïchiometry of binding, any associated conformational change and to identify promiscuous inhibitors. Hit ranking and clustering: Confirmed hit compounds are then ranked according to the various hit confirmation experiments.

Hit expansion

Following hit confirmation, several compound clusters will be chosen according to their characteristics in the previously defined tests. An Ideal compound cluster will have compound members that exhibit a high affinity towards the target (less than 1 μM). Moderate molecular weight and lipophilicity (usually measured as cLogP). Affinity, molecular weight and lipophilicity can be linked in single parameter such as ligand efficiency and lipophilic efficiency to assess drug likeness

- show chemical tractability
- be free of Intellectual property
- not interfere with the P450 enzymes nor with the P-glycoproteins
- not bind to human serum albumin
- be soluble in water (above 100 μM)
- be stable
- have a good druglikeness
- exhibit cell membrane permeability
- show significant biological activity in a cellular assay
- not exhibit cytotoxicity
- not be metabolized rapidly
- show selectivity versus other related targets

The project team will usually select between three and six compound series to be further explored. The next step will allow to test analogous compounds to define Quantitative structure-activity relationship (QSAR). Analogs can be quickly selected from an internal library or purchased from commercially available sources. Medicinal chemists will also start synthesizing related compounds using different methods such as combinatorial chemistry, high-throughput chemistry or more classical organic chemistry synthesis.

Lead optimization phase

The objective of this drug discovery phase is to synthesize lead compounds, new analogs with improved potency, reduced off-target activities, and physiochemical/metabolic properties suggestive of reasonable in vivo pharmacokinetics. This optimization is accomplished through chemical modification of the hit structure, with modifications chosen by employing structure-activity analysis (SAR) as well as structure-based design if structural information about the target is available [7].

Fragment-based lead discovery

Fragment-based lead discovery (FBLD) also known as fragment-based drug discovery (FBDD) is a method used for finding lead compounds as part of the drug discovery process. It is based on identifying small chemical fragments, which may bind only weakly to the biological target, and then growing them or combining them to produce a lead with a higher affinity. FBLD can be compared with high-throughput screening (HTS). In HTS, libraries with up to millions of compounds, with molecular

weights of around 500 Da, are screened, and nanomolar binding affinities are sought. In contrast, in the early phase of FBLD, libraries with a few thousand compounds with molecular weights of around 200 Da may be screened, and millimolar affinities can be considered useful [8].

Library design

In analogy to the rule of five, it has been proposed that ideal fragments should follow the 'rule of three' (molecular weight < 300, ClogP < 3, the number of hydrogen bond donors and acceptors each should be < 3 and the number of rotatable bonds should be < 3). Since the fragments have relatively low affinity for their targets, they must have high water solubility so that they can be screened at higher concentrations.

Advantages over traditional libraries

Advantages of screening low molecular weight fragment based libraries over traditional higher molecular weight chemical libraries are several. These include:
More hydrophilic hits in which hydrogen bonding is more likely to contribute to affinity (enthalpically driven binding). It is generally much easier to increase affinity by adding hydrophobic groups (entropically driven binding), starting with a hydrophilic ligand increases the chances that the final optimized ligand will not be too hydrophobic (log P < 5).

Higher ligand efficiency so that the final optimized ligand will more likely be relatively low in molecular weight (MW < 500). Since two to three fragments in theory can be combined to form an optimized ligand, screening a fragment library of N compounds is equivalent to screening N2 - N3 compounds in a traditional library. Fragments are less likely to contain sterically blocking groups that interfere with an otherwise favorable ligand-protein interaction, increasing the combinatorial advantage of a fragment library even further [9].

High-content screening

High content screening (HCS) also known as high content analysis (HCA), or visual screening is a method that is used in biological research and drug discovery to identify substances such as small molecules, peptides, or RNAi that alter the phenotype of a cell in a desired manner. Hence high content screening is a type of phenotypic screen conducted in cells. Phenotypic changes may include increases or decreases in the production of cellular products such as proteins and/or changes in the morphology (visual appearance) of the cell. High content screening includes any method used to analyze whole cells or components of cells with simultaneous readout of several parameters. Hence the name high content screening.

In high content screening, cells are first incubated with the substance and after a period of time, structures and molecular components of the cells are analyzed. The most common analysis involves labeling proteins with fluorescent tags, and finally changes in cell phenotype are measured using automated image analysis. Through the use of fluorescent tags with different absorption and emission maxima, it is possible to measure several different cell components in parallel. Furthermore, the imagining is able to detect changes at a subcellular level (e.g., cytoplasm vs. nucleus vs. other organelles). Therefore a large number of data points can be collected per cell. In addition to fluorescent labeling, various label free assays have been used in high content screening.

General principles

High-content screening (HCS) in cell-based systems uses living cells as tools in biological research to elucidate the workings of normal and diseased cells. HCS is also used to discover and optimizes new drug candidates. High content screening is a combination of modern cell biology, with all its molecular tools, with automated high resolution microscopy and robotic handling. Cells are first exposed to chemicals or RNAi reagents. Changes in cell morphology are then detected using image analysis. Changes in the amounts of proteins synthesized by cells are measured using a variety of techniques such as the green fluorescent proteins fused to endogenous proteins, or by fluorescent antibodies.

The technology may be used to determine whether a potential drug is disease modifying. For example, in humans G-protein coupled receptors (GPCRs) are a large family of around 880 cell surface proteins that transduce extra-cellular changes in the environment into a cell response, like triggering an increase in blood pressure because of the release of a regulatory hormone into the blood stream. Activation of these GPCRs can involve their entry into cells and when this can be visualised it can be the basis of a systematic analysis of receptor function through chemical genetics, systematic genome wide screening or physiological manipulation.

At a cellular level, parallel acquisition of data on different cell properties, for example activity of signal transduction cascades and cytoskeleton integrity is the main advantage of this method in comparison to the faster but less detailed high throughput screening. While HCS is slower, the wealth of acquired data allows a more profound understanding of drug effects. Automated image based screening permits the identification of small compounds altering cellular phenotypes and is of interest for the discovery of new pharmaceuticals and new cell biological tools for modifying cell function. The selection of molecules based on a cellular phenotype does not require a priori knowledge of the biochemical targets that are affected by compounds. However the identification of the biological target will make subsequent preclinical optimization and clinical development of the compound hit significantly easier. Given the increase in the use of phenotypic/visual screening as a cell biological tool,

methods are required that permit systematic biochemical target identification if these molecules are to be of broad use. Target identification has been defined as the rate limiting step in chemical genetics/high-content screening [10].

Instrumentation
An automated confocal image reader

High-content screening technology is mainly based on automated digital microscopy and flow cytometry, in combination with IT-systems for the analysis and storage of the data. High-content or visual biology technology has two purposes, first to acquire spatially or temporally resolved information on an event and second to automatically quantify it. Spatially resolved instruments are typically automated microscopes, and temporal resolution still requires some form of fluorescence measurement in most cases. This means that a lot of HCS instruments are (fluorescence) microscopes that are connected to some form of image analysis package. These take care of all the steps in taking fluorescent images of cells and provide rapid, automated and unbiased assessment of experiments.

The instruments on the market can be divided on the basis of price, footprint and the ethereal design qualities of the box they come in - but the most incisive difference is whether the instruments are optical confocal or not. Confocal imaging summarizes as imaging/resolving a thin slice through an object and rejecting out of focus light that comes from outside this slide. This gives higher image signal to noise and higher resolution than the more commonly applied epi-fluorescence microscopy. For many biological assays, confocal imaging is not ideal (e.g. phototoxicity issues or the need for a larger focal depth etc.). What all instruments share is the ability to take, store and interpret images automatically and most integrate into large robotic cell/medium handling platforms.

Applications

This technology allows a (very) large number of experiments to be performed, allowing explorative screening. Cell-based systems are mainly used in chemical genetics where large, diverse small molecule collections are systematically tested for their effect on cellular model systems. Novel drugs can be found using screens of tens of thousands of molecules, and these have promise for the future of drug development. Beyond drug discovery, chemical genetics is aimed at functionalizing the genome by identifying small molecules that acts on most of the 21,000 gene products in a cell. High-content technology will be part of this effort which could provide useful tools for learning where and when proteins act by knocking them out chemically. This would be most useful for gene where knockout mice (missing one or several genes) cannot be made because the protein is required for development, growth or otherwise lethal when it is not there. Chemical knock out could address how and where these genes work [11].

CONCLUSION

Further the technology is used in combination with RNAi to identify sets of genes involved in specific mechanisms, for example cell division. Here, libraries of RNAis, covering a whole set of predicted genes inside the target organisms genome can be used to identify relevant subsets, facilitating the annotation of genes for which no clear role has been established beforehand. The large datasets produced by automated cell biology contain spatially resolved, quantitative data which can be used for building for systems level models and simulations of how cells and organisms function. Systems biology models of cell function would permit prediction of why, where and how the cell responds to external changes, growth and disease.

REFERENCES

1. Steven M. Paul, Daniel S. Mytelka, Christopher T. Dunwiddie, Charles C. Persinger, Bernard H. Munos, Stacy R. Lindborg & Aaron L. Schacht. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, 9(3), 2010, 203–214.
2. Hopkins AL, Groom CR and Alexander A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today*, 9(10), 2004, 430–431.
3. Ryckmans T. et al. Rapid assessment of a novel series of selective CB2 agonists using parallel synthesis protocols: A Lipophilic Efficiency (LipE) analysis. *Bioorg. Med. Chem. Lett*, 19(15), 2009, 4406–4409.
4. Leeson PD. et al. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Disc*, 6(11), 2007, 881–890.
5. Feher M, Schmidt JM. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci*, 43(1), 2003, 218–27.
6. Newman DJ, Cragg GM. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod*, 70(3), 2007, 461–77.
7. Von Nussbaum F, Brands M, Hinzen B, Weigand S, Häbich D. Antibacterial natural products in medicinal chemistry--exodus or revival?. *Angew. Chem. Int. Ed. Engl*, 45(31), 2006, 5072–129.
8. John Faulkner D, Newman DJ, Cragg GM. Investigations of the marine flora and fauna of the Islands of Palau. *Nat Prod Rep*, 21(1), 2004, 50–76.

9. Gasparri F. An overview of cell phenotypes in HCS: limitations and advantages. *Expert Opinion on Drug Discovery*, 4(6), 2009, 643–657.
10. Burdine L, Kodadek T. Target identification in chemical genetics: *Chem. Biol*, 11(5), 2004, 593-7.
11. Eggert US, Mitchison TJ. Small molecule screening by imaging. *Curr Opin Chem Biol*, 10(3), 2006, 232–7.