Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
# The Molecular Libraries Screening Center Network (MLSCN): Identifying Chemical Probes of Biological Systems

**Donna M. Huryn** and **Nicholas D. P. Cosford**

## Contents

1. Introduction 402
   1.1 Goals and overview of the MLSCN 402
   1.2 Description of MLSCN Centers 403
   1.3 PubChem 407
2. Role of Chemistry in the MLSCN 408
   2.1 Molecular Libraries Small Molecule Repository 408
   2.2 Chemistry cores 408
   2.3 Pilot Scale Libraries and Chemical Methodologies and Library Development Centers 408
3. Novel Probes from the MLSCN 409
   3.1 MKP-1 inhibitors (PMLSC) 409
   3.2 BID inhibitors by NMR-based screening (SDCCG) 410
   3.3 Cathepsin B alternate substrate (PCMD) 410
   3.4 Pyruvate kinase probes (NCGC) 410
   3.5 Glucocerebrosidase inhibitor probes (NCGC) 411
   3.6 S1P1 antagonist probe (Scripps) 412
   3.7 SARS CoV inhibitors (SRMLSC) 412
   3.8 Pantothenate synthetase inhibitors (SRMLSC) 413
   3.9 GPCR30 antagonist (New Mexico MLSC) 414
4. Future Outlook and Conclusions 414

References 415

* Pittsburgh Molecular Libraries Screening Center, University of Pittsburgh, Pittsburgh, PA 15260 and Penn Center for Molecular Discovery, University of Pennsylvania, Philadelphia, PA 19104, USA
** San Diego Center for Chemical Genomics, Burnham Institute for Medical Research, 10901 N. Torrey Pines Road, La Jolla, CA 92037, USA

---

*Pittsburgh Molecular Libraries Screening Center, University of Pittsburgh, Pittsburgh, PA 15260 and Penn Center for Molecular Discovery, University of Pennsylvania, Philadelphia, PA 19104, USA

**San Diego Center for Chemical Genomics, Burnham Institute for Medical Research, 10901 N. Torrey Pines Road, La Jolla, CA 92037, USA

Annual Reports in Medicinal Chemistry, Volume 42
ISSN 0065-7743, DOI 10.1016/S0065-7743(07)42026-7
© 2007 Elsevier Inc. All rights reserved.
1. INTRODUCTION

1.1 Goals and overview of the MLSCN

The NIH Molecular Libraries Screening Center Network (MLSCN) is a subset of the Molecular Libraries Initiative (MLI) component of the NIH Roadmap for Medical Research [1,2]. It consists of a consortium of 10 centers, each having expertise in assay development, high-throughput screening (HTS), chemistry and informatics. Using a centralized screening library of approximately 100,000 small molecules and assays from the research community, scientists at each center optimize assays, carry out high-throughput screens, and deposit the results into PubChem (vida infra). Based on the data from those assays, chemists at each center further optimize the initial hits in order to develop unique, small molecule probes of biological systems [3]. Informaticists contribute to data handling and analysis throughout this process. The chemical probes developed are available to researchers (both public and private sectors) via data deposition into PubChem, and in the future, via access to samples and synthetic protocols to prepare the probes. These efforts support the ultimate goals of the MLSCN and the MLI, which is to “expand the availability, flexibility, and use of small-molecule chemical probes for basic research” [1].

A number of aspects of the MLSCN make this initiative unique from other academic screening centers [4,5], as well as from screening and lead optimization activities being undertaken at pharmaceutical and biotechnology companies. First, all researchers (public and private sector) have access to the screening centers through the NIH X01 and R03 funding mechanisms [6]. Second, due to the diverse source of assays and the wide expertise available within the MSLCN, specific biological systems investigated and screened will include: (a) “high risk” targets, that is proteins or biological systems whose function is unknown; (b) targets implicated in orphan diseases or diseases not typically addressed by the private sector; (c) novel or uncommon assay systems (e.g., zebrafish, high content screening); and (d) “non-druggable” targets, such as inhibitors of aggregation and protein–protein interactions. Third, the small molecule screening library contains structures not typically found in commercial collections or those housed in pharmaceutical companies. Sources of these unique structures include natural products and novel compound libraries prepared by academic investigators through the Pilot Scale Libraries (PSL) granting mechanisms [7], compounds generated by the Centers for Chemical Methodologies and Library Development (CMLD) (vida infra) [8] and those obtained through solicitation by the NIH [9]. Fourth, as the goal of the MLSCN is to develop selective chemical probes and small molecule tools that will interrogate novel biochemical pathways, the criteria for an acceptable class of molecules is broader for the MLSCN than for those involved in drug discovery and development. Therefore, chemical probes are not subject to the same constraints on physical properties, functional groups or metabolic profiles that are common in the pharmaceutical industry, and necessary for successful clinical candidates. An example of a compound that would fit the definition of a valuable chemical probe, but which would not...
adhere to the commonly prescribed criteria of “drug-like” is a staurosporine-derived ruthenium complex, shown below, which is a selective, sub-nanomolar inhibitor of the kinase, Pim1 [10]. Fifth, and particularly important for this review, is the inclusion of integral medicinal chemistry within each MLSCN Center, that allows the network to produce chemical probes with particular properties, rather than simply identifying apparent activities from the screening collection. Finally, unlike all other screening efforts in both industry and academia, all data are available in PubChem with no delay in publication. In addition to HTS protocols and primary screening results, secondary assay data, “profiling data” (e.g., aggregation evaluation, Cytochrome P450 inhibition, spectroscopic profiling, solubility measurements), follow-up compound libraries and their associated biological data, and synthetic protocols are accessible.

\[
\begin{align*}
\text{Pim1 inhibitor (0.25 nM)}
\end{align*}
\]

1.2 Description of MLSCN Centers

1.2.1 NCGC
The NIH Chemical Genomics Center (NCGC) [11] is an ultrahigh-throughput screening (uHTS) and chemistry center that applies the tools of small molecule screening and discovery to develop chemical probes for the study of protein and cell functions. Using a process called quantitative high-throughput screening (qHTS), chemical libraries are screened at multiple concentrations (typically seven) to generate a concentration–response curve for each compound that covers a range of five orders of magnitude (typically 1 nM to 100 μM). qHTS comprehensively and efficiently characterizes biological activities of large chemical libraries to yield high-quality datasets for chemical probe development and compound profiling [12]. This process has been applied successfully to both cell-free and cell-based assays. The Kalypsys robotic system uses multimodal detectors, the ViewLux and Envision systems, and a plate-based laser cytometry system (Acumen Explorer) for high-capacity screening (100,000+wells/day) in the reagent-sparing 1536-well plate format. Assay detection capabilities include absorbance, luminescence, fluorescence resonance energy transfer (FRET), time-resolved FRET (TR-FRET), fluorescence polarization (FP), fluorescence intensity (FI), and AlphaScreen™, as well as cell-based imaging assays that employ
fluorescent proteins such as GFP (green fluorescence protein). The NCGC also develops new paradigms for screening, informatics, and chemical probe development that extend the application of small molecule technology to new areas of the genome. Collaborations with academic investigators worldwide, as well as pharmaceutical and biotechnology companies produces public domain data, thereby allowing sharing of best practices to enable both chemical genomics and downstream drug development. The NCGC accepts chemical libraries for screening from academic and industrial investigators, and produces its own focused libraries for specific projects that are added to the screening collection. With its miniaturized qHTS process, as little as 0.1–0.5 mg of compound will support several years of screening against hundreds of diverse biological assays. The NCGC is part of the intramural NIH program of the National Human Genome Research Institute.

1.2.2 PCMD
The Penn Center for Molecular Discovery (PCMD) [13] approaches the high volume screening challenge with unique capabilities. A key technology for the Center is the ability to print thousands of molecules on a glass surface the size of a business card, and then rapidly test these molecules against proteases and other enzymes purified from human or animal cells, bacteria, parasites, insects, or viruses [14]. Scientists at the Penn Center are also able to test compounds in thousands of miniature wells each containing a millimeter-sized *Danio rerio* (zebrafish), an unlikely organism that has proven its worth in studies of heart and nerve function, as well as in cancer biology, because the transparent fish is easily imaged. The PCMD, based in Philadelphia, is surrounded by local HTS industrial labs including those of Merck, GSK, Wyeth, and Johnson & Johnson. These industrial connections help transfer HTS skills into the university environment of this center.

1.2.3 Emory
The Emory Chemical Biology Center in the MLSCN has the capability to adapt and optimize all target-based and phenotypic assays selected by the MLSCN, but has identified protein–protein interactions for small molecule probe discovery as the Center’s theme. With two general screening platforms established, the Center is able to perform both HTS and high content screening (HCS) using a variety of *in vitro* biochemical assays, cell reporter assays, and cell phenotype-based assays. In particular, this center is experienced in assays for monitoring protein–protein interactions and enzyme activities with fluorescence-based assays, including FI, FP, and FRET. Examples of assays that are within this center’s capacity include protein–protein interactions (FI, FP, FRET, AlphaScreen™), enzyme assays (FI, FP, and other coupled assays), receptor–ligand interaction assays (FI, FRET, Ca²⁺ imaging), reporter assays (luciferase, GFP, etc.), viability assays and protein translocation assays (e.g., receptor internalization and membrane and nuclear localization).
1.2.4 PMLSC
In common with the other MLSCN centers, the overall goal of the Pittsburgh Molecular Libraries Screening Center (PMLSC) [15] is to provide the scientific community access to a facility that is designed to optimize, validate, and implement assays for HTS based on optical-based detection methods to identify chemical probes, and to deposit these data into the PubChem database. The PMLSC has been designed for maximum flexibility with regard to target classes and assay formats. They develop and implement cell based, biochemical and model organism \(\text{Danio rerio}\) (zebrafish) and \(\text{Drosophila melanogaster}\) (fly) assays preferentially in the 384-well plate format. Existing optical-based detection capabilities include automated HCS imaging platforms, absorbance, FI, FP, time-resolved fluorescence (TRF), FRET, and luminescence. The PMLSC also examines the structure–activity relationships of active small molecules and synthesizes probe molecules that demonstrate significant potency and target selectivity.

1.2.5 SRMLSC
The Southern Research Molecular Libraries Screening Center (SRMLSC) is based at Southern Research Institute (SRI), where more than 20 anti-cancer agents have been discovered and entered into clinical trials, six of which received FDA approval and proceeded to market. The SRMLSC brings extensive drug discovery and development expertise to the network, especially in the areas of cancer, neurological diseases/CNS disorders, and infectious disease (HIV, hepatitis, TB, and emerging pathogens including influenza, H5N1 Avian flu, West Nile virus, and SARS coronavirus). The screening center has broad capabilities to implement any cellular, molecular, or target-based assay including those which require BSL-3 containment, and optimizes or miniaturized them as necessary. The HTS facility is equipped with state-of-the-art instrumentation to screen in up to 1536-well plate format, including two ORCA robotic rails, multiple plate readers, and two Biomek FX liquid handlers, a BioRaptr, and an Echo 550 for nanoliter volume dispensing. In addition, the Center uses a high-speed automated Evotec Opera confocal microscope for high-throughput imaging assays. Data analysis is performed by scientists with expertise in molecular modeling, predictive algorithms, and QSAR analysis, using a robust assortment of chemoinformatics software packages.

1.2.6 SDCCG
The San Diego Center for Chemical Genomics (SDCCG) [16], located in the biotech-rich heart of La Jolla, California, has broad expertise in biochemistry and the ability to run almost any assay type. Specific biological themes include targets involved in regulating cell death, using a variety of biochemical and cell-based assays, with particular emphasis on kinases, phosphatases and proteases. Another area of expertise is in phenotypic assays for stem cell differentiation, using fluorescent reporters. Two main technological themes are incorporated into the SDCCG. First, the Center has special expertise in high-throughput microscopy as a tool for performing high-content, cell-based screens where cellular phenotypes drive compound selection in an unbiased manner. Second, the
Center is unique across the network in having the capability to perform NMR-based small-molecule screening and optimization. NMR-based methods are exceptionally valuable when investigating molecular targets that are not easily tractable by other methods, such as protein–protein interactions and protein targets that cannot be formatted for the classical HTS environment.

1.2.7 Scripps
The Scripps Research Institute Molecular Screening Center [17] spans the Scripps campuses in La Jolla and West Palm Beach, Florida. Scripps has brought together an integrated combination of infrastructure, people and technologies that can support the identification of proof-of-concept small molecules in the academic setting. These small molecules comprise chemical probes of adequate potency, selectivity, physical properties and stability to show robust activities in cell-based assays and in vivo, allowing pre-competitive advancement to fields of breaking biology. The center is equipped with a fully automated Kalypsys screening system, with plate hotels and incubators, 200 nL to 20 μL volume dispensing 1536-well aspiration, and 1536 pintool heads for compound delivery for uHTS screening of larger compound decks. Plate readers and detectors include the ViewLux CCD-based plate reader and the EnVision multimode detector equipped with Alpha Screen™. Assay formats include TRF, FP, FI, FRET, luminescence, and absorbance. This center is vertically integrated with enterprise-scale data management and chemoinformatics, high throughput LC/MS for compound quality assurance and rodent pharmacokinetics, and facilities for downstream synthetic follow-up by modular, library or linear chemical approaches. The goal of the Scripps Center is the rapid, collaborative publication of interesting compounds that advance the understanding of biological problems, or illuminate new nodal control points in physiology by short-term chemical perturbation.

1.2.8 Columbia
The strength and experience of the MLSCN Center at Columbia University are in cell biology, high content/high-resolution automated cellular imaging and image analysis, and phenotypic assay design and implementation. The main imaging platform of this center is the INCell Analyzer 3000 (GE Healthcare), a state of the art high throughput cell imaging system. The INCell instrument uses three laser lines for excitation: a Krypton laser (647 nm) and an Argon laser (364 and 488 nm). Three fluorescence channels can be recorded by three independent high-speed 12-bit CCD cameras, and emitted light in the wavelength range from 420 to 720 nm can be captured. Connected to a Kendro Plate Hotel and a Mitsubishi robotic arm, the system can image and analyze 222 plates (96/384 well) without supervision. Depending on the specific assay, up to 50,000 wells can be processed per day. A whole array of different image analysis modules is available, and analysis is performed at high speed on the fly. This imaging system enables the center to screen and analyze a very broad variety of assays monitoring a wide spectrum of biological processes.
1.2.9 New Mexico MLSC
The New Mexico Molecular Libraries Screening Center (New Mexico MLSC) [18] has developed innovative flow cytometry tools for discovery research that enable homogeneous analysis of ligand binding and protein–protein interactions, high throughput sample handling, high content analysis, and real-time measurements of cell response. Using their novel HyperCyt$^\text{®}$ screening technology, virtually any molecular assembly or cell response can be displayed in a HTS format compatible with flow cytometry, and assessing both cellular and molecular activities of small molecules is possible. Moreover, by creating a suspension array of particles, assays and responses can be highly multiplexed or performed on complex cell populations without loss of throughput. It is likely that no single competing technology offers the versatility of flow cytometry for MLI screening or has the potential of being available to such a large number of laboratories that house flow cytometers (20,000 world-wide). The Center brings together expertise that spans biomedical, biophysical, chemical, computational, instrumentation, and engineering disciplines, and is particularly interested in enhancing the overall discovery process through the integration of physical screening and computational tools that include virtual screening, chemoinformatics, and data mining.

1.2.10 Vanderbilt
The goal of the Vanderbilt Screening Center for G-protein Coupled Receptors (GPCRs), Ion Channels, and Transporters [19] is to enable investigators to discover and develop a new generation of small molecule probes to promote our understanding of physiological and disease processes, with a particular emphasis on the structure and function of GPCRs, ion channels, and transporters. Measurements for biochemical, cellular and cell-free assays are made using a wide variety of commercially available and novel technologies. The suite of detection modalities includes two Hamamatsu FDSS kinetic imaging plate readers. These instruments are capable of collecting data from all wells of 96 or 384 plates simultaneously, and during integrated reagent addition at up to 10 frames per second over wavelengths from UV to far red with dual excitation, emission, and fluorescence polarization modes. The FDSS also supports ultra low-light detection for aequorin and other kinetic/flash luminescence formats. Additionally, the Vanderbilt Center supports high-content screening through the use of the BlueShift Isocyte, a laser scanning fluorimeter that generates two-dimensional anisotropy data. This combination of capabilities paired with robust automation provides tremendous flexibility for measuring the action of a test compound on a wide range of targets. The investment in infrastructure, the combination of basic and industrial research expertise, the dedication to translational and chemical biology, and the establishment and maintenance of a highly collaborative environment make the Vanderbilt Center well suited to support the MLSCN.

1.3 PubChem
PubChem is a comprehensive, publicly accessible database developed by the National Center for Biotechnology Information at the National Library of Medicine
that contains information on the biological activities of small molecules [20]. As of March 2007, PubChem contained more than 15 million records, 10 million unique structures, and data from over 400 assays. The database is linked to other Entrez databases such as PubMed and PubChem Central [21]. All data (e.g., assay results, secondary assays, structures of compounds synthesized) generated within the MLSCN is deposited into PubChem. Access to this range of data on a large library of diverse compounds has enormous potential for use by the Informatics community for the development of computational models, pharmacophore models, and other algorithms to predict biological activities and properties of small molecules.

2. ROLE OF CHEMISTRY IN THE MLSCN

Two distinct aspects of the MSLCN require participation by, and input from chemists: first, synthetic chemistry is a source of compounds within the screening library, and therefore of the assay hits; and second, expertise in synthetic and medicinal chemistry is required to optimize the hits into usable probes of biological systems.

2.1 Molecular Libraries Small Molecule Repository

The Molecular Libraries Small Molecule Repository (MLSMR), managed by Biofocus/DPI (under contract from NIH) collects, maintains, and distributes compound samples to the MLSCN [22]. Compounds, collected from a variety of commercial and non-commercial sources, such as PSL grantees, CMLDs (vida infra), and other academic investigators, are filtered through integrity, purity, and physical property criteria before being accepted into the MLSMR. Once accepted, samples are acquired and distributed to each MLSCN Center, and structures and associated data are deposited into PubChem. The approximately 100,000 compounds that currently comprise the MLSMR will grow to 300,000 during 2007, with a goal of reaching up to 1 million compounds over the next decade [23].

2.2 Chemistry cores

Each screening center has medicinal and synthetic chemistry expertise in order to optimize hits identified from HTS campaigns and develop them into chemical probes. Specific capabilities vary, however typical strategies employed include parallel synthesis, computational and informatics analysis, and analytical capabilities such as LC/MS techniques. The structures of novel compounds that are prepared, their synthetic protocols, analytical data and biological data are all available, and samples of final probes developed are deposited into the MLSMR. A Working Group comprised of chemists from each center meets regularly to share information, best practices, and insure optimal use of resources.

2.3 Pilot Scale Libraries and Chemical Methodologies and Library Development Centers

Another key component of the Molecular Libraries Initiative is the development of novel technologies for generating chemical diversity, the application of those
technologies to compound library synthesis or generation, and the deposition of those compounds into the MLSMR. Towards this end, NIH has established two funding opportunities. The objective of the Pilot Scale Library granting mechanism is to generate chemical libraries in order “to increase the diversity and the uniqueness” of compounds contained within the MLSMR [7]. Both synthetic and natural product-derived compounds are included in this initiative. Centers for Chemical Methodologies and Library Development [8] are multi-investigator research programs that develop state-of-the-art synthetic methodologies for the design and production of chemically diverse, small molecule libraries. Both initiatives contribute compounds to the MLSMR.

3. NOVEL PROBES FROM THE MLSCN

3.1 MKP-1 inhibitors (PMLSC)

The Pittsburgh Molecular Libraries Screening Center reported the identification of SID 3717140 as an inhibitor of Mitogen-activated Protein Kinase Phosphatase 1 (MKP-1) [24]. This initial hit exhibited only modest potency (IC$_{50}$ = 19.2 μM), however it appeared to display some selectivity against other phosphatases, and excellent selectivity against the 52 other targets tested [25]. A small library of compounds was designed and prepared in an effort to identify compounds with improved potency. Towards that end, several new uracil-based compounds, such as SID 14715524, exhibited improvements in potency. MKP-1 is a dual-specificity phosphatase involved in a number of processes related to cell proliferation. The availability of potent, selective and cell permeable probes would help enable a thorough understanding of the role this enzyme plays in cell cycle, signal transduction, oncogenesis, and apoptosis. While small molecule inhibitors of MKP-1 have been previously reported [26], they have been hampered by low in vitro potency, lack of cellular activity, and poor selectivity. As such, these uracil quinolines from the PMLSC represent a novel structural class which, based on their promising physicochemical properties, may provide an improvement over those inhibitors previously reported.
3.2 BID inhibitors by NMR-based screening (SDCCG)

Bcl-2 family proteins play a crucial role in tissue homeostasis and apoptosis (programmed cell death). The BH3-interacting domain death agonist (BID) is a proapoptotic member of the Bcl-2 family, promoting cell death when activated by caspase-8, which cleaves BID to its truncated active form, tBID. NMR-based screening of a library composed of 300 fragments followed by SAR optimization by interligand NOE led to the identification of two chemical fragments that bind on the surface of BID. Covalent linkage of the two fragments provided high-affinity bidentate derivatives such as BI–11A7 [27,28]. In vitro and cellular assays showed that these compounds prevent tBID translocation to the mitochondrial membrane and the subsequent release of proapoptotic stimuli, and inhibit neuronal apoptosis in the low micromolar range. These compounds may lead to therapeutic agents with the potential to treat disorders associated with BID activation including neurodegenerative diseases, cerebral ischemia, and brain trauma.

\[
\begin{align*}
&\text{H}_2\text{N} & \text{N} & \text{O} & \text{H} \\
&\text{H}_2\text{N} & \text{N} & \text{O} & \text{S} \\
&\text{H}_2\text{N} & \text{N} & \text{O} & \text{S} \\
\end{align*}
\]

BI-11A7

3.3 Cathepsin B alternate substrate (PCMD)

Screening of over 66,000 compounds from the MLSMR by scientists at the PCMD for inhibitors of Cathepsin B resulted in the identification and characterization of an alternate substrate, SID 16952359 [29]. This study also describes issues relating to the nucleophilicity of dithiothreitol (DTT) and cysteine, reductants frequently used in HTS protocols, and the potential for reactivity with electrophilic sites of probe molecules.

\[
\begin{align*}
&\text{H}_2\text{N} & \text{O} & \text{S} & \text{S} & \text{O} & \text{S} & \text{O} \\
\end{align*}
\]

SID 16952359

3.4 Pyruvate kinase probes (NCGC)

The development of quantitative high-throughput screening (qHTS) paradigms that provide concentration–response curves for large chemical libraries in a
single experiment is a major focus of the NIH Chemical Genomics Center [12]. This strategy was applied to a screen for inhibitors of the enzyme pyruvate kinase, and allowed SAR development directly from primary screening data, and rapid analysis and triage of active clusters. From this analysis, a class of cyano-oxazole inhibitors, exemplified by SID 862236, was identified that exhibited activity in the nanomolar range ($AC_{50} = 30 \text{nM}$). Importantly, structurally related analogs exhibited a range of potency from the nanomolar to inactive ranges. Activators were also identified in the same HTS experiment: SID 3712493 activated pyruvate kinase at an $AC_{50}$ concentration of 600 nM [30].

$$\text{SID 862236}$$

$$\text{SID 3712493}$$

3.5 Glucocerebrosidase inhibitor probes (NCGC)

Gaucher’s Disease is an inherited disorder characterized by deficiencies of glucocerebrosidase activity. An assay to identify small molecule inhibitors of glucocerebrosidase was developed [31], and three probes were identified. NCGC00092410 was identified by testing a series of purchased analogs of an initial hit. SID 4264637 and SID 847960 were members of the initial screening library [32]. Several of these probes have been shown to restore glucocerebrosidase activity in cultured cells from Gaucher patients, a result consistent with the correction of trafficking of misfolded glucocerebrosidase.

$$\text{NCGC00092410}$$

$$\text{SID 4264637}$$
Sphingosine 1-phosphate (S1P) regulates vascular barrier and lymphoid development, as well as lymphocyte egress from lymphoid organs, by activating high-affinity S1P\textsubscript{1} receptors. Based on phosphate esters (such as the structure below where $X = O$), the reversible S1P\textsubscript{1} antagonist ($X = \text{CH}_2$) was designed to provide a non-reactive chemical probe with \textit{in vivo} activity [33]. This compound was used to gain mechanistic insights into S1P systems organization not accessible through genetic manipulations and to investigate their potential for therapeutic modulation. Vascular (but not airway) administration of the preferred R enantiomer of this compound induced the loss of capillary integrity in mouse skin and lung, but did not affect the number of constitutive blood lymphocytes. Instead, alteration of lymphocyte trafficking and phenotype required supraphysiological elevation of S1P\textsubscript{1} tone and was reversed by the antagonist. \textit{In vivo} two-photon imaging of lymph nodes confirmed requirements for obligate agonism, and the data were consistent with the presence of a stromal barrier mechanism for gating lymphocyte egress. Chemical modulation revealed differences in S1P-S1P\textsubscript{1} ‘set points’ among tissues and highlights both mechanistic advantages (lymphocyte sequestration) and risks (pulmonary edema) of therapeutic intervention.

3.7 SARS CoV inhibitors (SRMLSC)

Researchers at SRMLSC recently developed a HTS that allowed the identification of potential inhibitors of the severe acute respiratory syndrome coronavirus (SARS CoV) from large compound libraries [34]. The luminescent-based assay, which measured the inhibition of SARS CoV-induced cytopathic effects (CPE) in Vero E6 cells, was validated with two different diversity sets of compounds against the SARS CoV. The hit rate for both libraries was approximately 0.01%.
The validated HTS assay was then employed to screen a 100,000-compound library against SARS CoV. The hit rate for the library in a single-dose format was determined to be approximately 0.8%. Screening of the three libraries resulted in the identification of several novel compounds that effectively inhibited the CPE of SARS CoV in vitro. Three hit compounds, shown below, were identified as promising lead candidates for further evaluation.

3.8 Pantothenate synthetase inhibitors (SRMLSC)

The team at SRMLSC also recently developed a screen for pantothenate synthetase (PS). PS (EC 6.3.2.1) is encoded by the panC gene and catalyzes the essential adenosine triphosphate (ATP)-dependent condensation of D-pantoate and β-alanine to form pantothenate in bacteria, yeast, and plants. Pantothenate is a key precursor for the biosynthesis of coenzyme A (CoA) and acyl carrier protein (ACP). Because the enzyme is absent in mammals, and both CoA and ACP are essential cofactors for bacterial growth, PS is an attractive chemotherapeutic target. An automated high-throughput screen was developed to identify drugs that inhibit Mycobacterium tuberculosis PS. The activity of PS was measured spectrophotometrically through an enzymatic cascade involving myokinase, pyruvate kinase, and lactate dehydrogenase. The rate of PS ATP utilization was quantitated by the reduction of absorbance due to the oxidation of NADH to NAD+ by lactate dehydrogenase, which allowed for an internal control to detect interference from compounds that absorb at 340 nm. This coupled enzymatic reaction was used to screen 4080 compounds in a 96-well format. This led to the discovery of a novel inhibitor of PS that exhibits potential as an antimicrobial agent [35].
3.9 GPCR30 antagonist (New Mexico MLSC)

Researchers at the New Mexico MLSC used a combination of virtual and biomolecular screening to discover a selective agonist of GPR30. Estrogen is a hormone critical in the development, normal physiology and pathophysiology of numerous human tissues. The effects of estrogen have traditionally been solely ascribed to estrogen receptor α (ERα) and more recently ERβ, members of the soluble nuclear ligand-activated family of transcription factors. However, it was recently shown that the GPR30 binds estrogen with high affinity and resides in the endoplasmic reticulum, where it activates multiple intracellular signaling pathways. To differentiate between the functions of ERα, ERβ and GPR30, the New Mexico MLSC team used a combination of virtual and biomolecular screening to isolate compounds that selectively bind to GPR30. Further studies led to the identification of the first GPR30-specific agonist, G-1 (shown below) capable of activating GPR30 in a complex environment of classical and new estrogen receptors [36].

![Structure of G-1](image_url)

4. FUTURE OUTLOOK AND CONCLUSIONS

The completion of the human genome project, a coordinated effort between government, academia, and industry, has prompted a vast expanse of medical research focused on the understanding of the fundamental causes of human disease. The Molecular Libraries and Imaging initiative is a natural extension of that groundbreaking effort. By providing access to assays, HTS capabilities, small molecule libraries and chemical optimization expertise, novel chemical probes are being developed that will allow the study of gene function, biochemical pathways, and cellular biology. It is also possible that through this initiative, starting points for new drugs, particularly of rare diseases, will be identified. The data deposited into PubChem via this effort will also serve as an unprecedented source of information for scientists in all biomedical disciplines. The availability of large datasets of biological activity on a common set of compounds should serve to stimulate advances in computational and predictive models of biological activities and chemical properties. This dataset should also expedite biomedical research through its use to evaluate selectivity, toxicity, and off-target activities of compounds. The value and success of this initiative may not be obvious or measurable for a number of years, but the adoption of some of the technologies
and approaches (e.g., HTS, hit optimization) typically available only inside the pharmaceutical industry should provide training, opportunities and inspiration to the wider scientific community.

REFERENCES

[1] C. P. Austin, L. S. Brady, T. R. Insel and F. S. Collins, Science, 2004, 306, 1138.
[2] J. S. Lazo, L. S. Brady and R. Dingledine, Mol. Pharmacol., 2007, 72, 1.
[3] See: http://nihroadmap.nih.gov/molecularlibraries/
[4] N. Tolliday, P. A. Clemons, P. Ferrariolo, A. N. Koehler, T. A. Lewis, X. Li, S. L. Schreiber, D. S. Gerhard and S. Eliasof, Cancer Res., 2006, 66, 8935.
[5] E. J. Gordon, Chem. Biol., 2007, 2, 9.
[6] http://grants.nih.gov/grants/guide/notice-files/NOT-RM-07-003.html.
[7] http://www.nigms.nih.gov/Initiatives/CMLD; For individual center web-sites see: http://cmdl.bu.edu/; http://www.cmdl.ku.edu/, http://www.broad.harvard.edu/chembio/cmld/index.html; http://ccc.chem.pitt.edu/
[9] http://grants.nih.gov/grants/notice-files/NOT-RM-07-005.html.
[10] H. Bregman, P. J. Carroll and E. Meggers, J. Am. Chem. Soc., 2006, 128, 877.
[11] http://www.ncgc.nih.gov/
[12] J. Inglese, D. S. Auld, A. Jadhav, R. L. Johnson, A. Simeonov, A. Yasgar, W. Zhang and C. P. Austin, Proc. Natl. Acad. Sci. USA, 2006, 103, 11473.
[13] http://www.seas.upenn.edu/~pcmd/
[14] D. Gosalia and S. Diamond, Proc. Natl. Acad. Sci. USA, 2003, 100, 8721.
[15] http://pmlsc.pitt.edu/default.asp
[16] http://sdccg.burnham.org/metadot/index.pl?iid = 2306&isa = Category
[17] http://molscreen.florida.scripps.edu/mlscn.html
[18] http://screening.health.unm.edu/
[19] http://www.vanderbilt.edu/MLSCN/Templates/index.htm
[20] http://pubchem.ncbi.nlm.nih.gov/help.html#PubChem_Overview
[21] D. L. Wheeler, T. Barrett, D. A. Benson, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. DiCuccio, R. Eđgar, S. Federhen, L. Y. Geer, Y. Kapustin, O. Khovayko, D. Landsman, D. J. Lipman, T. L. Madden, D. R. Maglott, J. Ostell, V. Miller, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, K. Sirotkin, A. Souvorov, G. Starchenko, R. L. Tatusov, T. A. Tatusova, L. Wagner and E. Yaschenko, Nucl. Acid Res., 2007, 35, D5.
[22] http://mlsmr.glpg.com/MLSMR_HomePage/index.html
[23] D. Livingston and T. Lease, SBS News 19 (2005) 1; (http://www.sbsonline.org/publications/news/issues/2005_12/index.php)
[24] D. A. Arnold, C. Foster, D. M. Huryn, J. S. Lazo, P. A. Johnston and P. Wipf, Chem. Biol. Drug Des., 2007, 69, 23.
[25] http://pubchem.ncbi.nlm.nih.gov/
[26] A. Vogt, K. A. Cooley, M. Brisson, M. G. Tarpey, P. Wipf and J. S. Lazo, Chem. Biol., 2003, 10, 733; A. Vogt, A. Tamewitz, J. Skoko, R. P. Sikorski, K. A. Guiliano and J. S. Lazo, J. Biol. Chem., 2005, 280, 19078; J. S. Lazo, R. Nunes, J. J. Skoko, R. E. Q. Oliveira, A. Vogt and P. Wipf, Bioorg. Med. Chem., 2002, 27, 5643.
[27] B. Becattini, C. Culumsee, M. Leone, D. Zhai, X. Zhang, K. J. Crowell, M. F. Rega, S. Landshamer, J. C. Reed, N. Plesnila and M. Pellecchia, Proc. Natl. Acad. Sci. USA, 2006, 103, 12602.
[28] B. Becattini, S. Sareth, D. Zhai, K. J. Crowell, M. Leone, J. C. Reed and M. Pellecchia, Chem. Biol., 2004, 11, 1107.
[29] M. C. Myers, A. D. Napper, N. Motlekar, P. P. Shah, C.-H. Chiu, M. P. Beavers, S. L. Diamond, D. M. Huryn and A. B. Smith, III, Bioorg. Med. Chem. Lett., 2007 (in press).
[30] http://ncgc.nih.gov/reports/PyruvateKinaseProbeReport2006-1128.pdf
[31] http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid = 360
[32] W. Zheng, J. Padia, D. J. Urban, A. Jadhav, O. Goker-Alpan, A. Simeonov, E. Goldin, D. Auld, M. E. LaMarca, J. Inglese, C. P. Austin and E. Sidransky, Proc. Natl. Acad. Sci., 2007, 104, 13192.

[33] M. G. Sanna, S.-K. Wang, P. J. Gonzalez-Cabrera, A. Don, D. Marsolais, M. P. Matheu, S. H. Wei, I. Parker, E. Jo, W.-C. Cheng, M. D. Cahalan, C.-H. Wong and H. Rosen, Nat. Chem. Biol., 2006, 2, 434.

[34] W. E. Sevenson, N. Shindo, M. Sosa, T. Fletcher, E. L. White, S. Ananthan and C. B. Jonsson, J. Biomol. Screen, 2007, 12, 33.

[35] E. L. White, K. Southworth, L. Ross, S. Cooley, R. B. Gill, M. I. Sosa, A. Manouvakhova, L. Rasmussen, C. Goulding, D. Eisenberg and T. M. Fletcher, J. Biomol. Screen, 2007, 12, 100.

[36] C. G. Bologa, C. M. Revankar, S. M. Young, B. S. Edwards, J. B. Arterburn, A. S. Kiselyov, M. A. Parker, S. E. Tkachenko, N. P. Savchuck, L. A. Sklar, T. I. Oprea and E. R. Prossnitz, Nat. Chem. Biol., 2006, 2, 207.