NIH Roadmap
Molecular Libraries and Imaging:

Opportunities for Investigators to Interact with the Molecular Libraries Screening Centers Network

Roadmap

Since 2002, NIH has begun a series of far-reaching initiatives, NIH Roadmap, that is a vision, strategic plan and guidelines for medical research in the 21st century.

Mission

New Pathways to Discovery sets out to build a better "toolbox" for medical research in the 21st century.

To empower the research community to use small molecule compounds in their research, whether as tools to perturb genes and pathways, as imaging probes in basic or clinical applications, or as starting points to the development of new therapeutics for human disease.

New Pathways to Discovery

- Building Blocks, Biological Pathways, and Networks
- Molecular Libraries and Imaging
- Structural Biology
- Bioinformatics and Computational Biology
- Nanomedicine

Research Teams of the Future

- High-Risk Research
  - NIH Director's Pioneer Award
- Interdisciplinary Research
- Public-Private Partnerships

Re-engineering the Clinical Research Enterprise

- Re-engineering the Clinical Research Enterprise

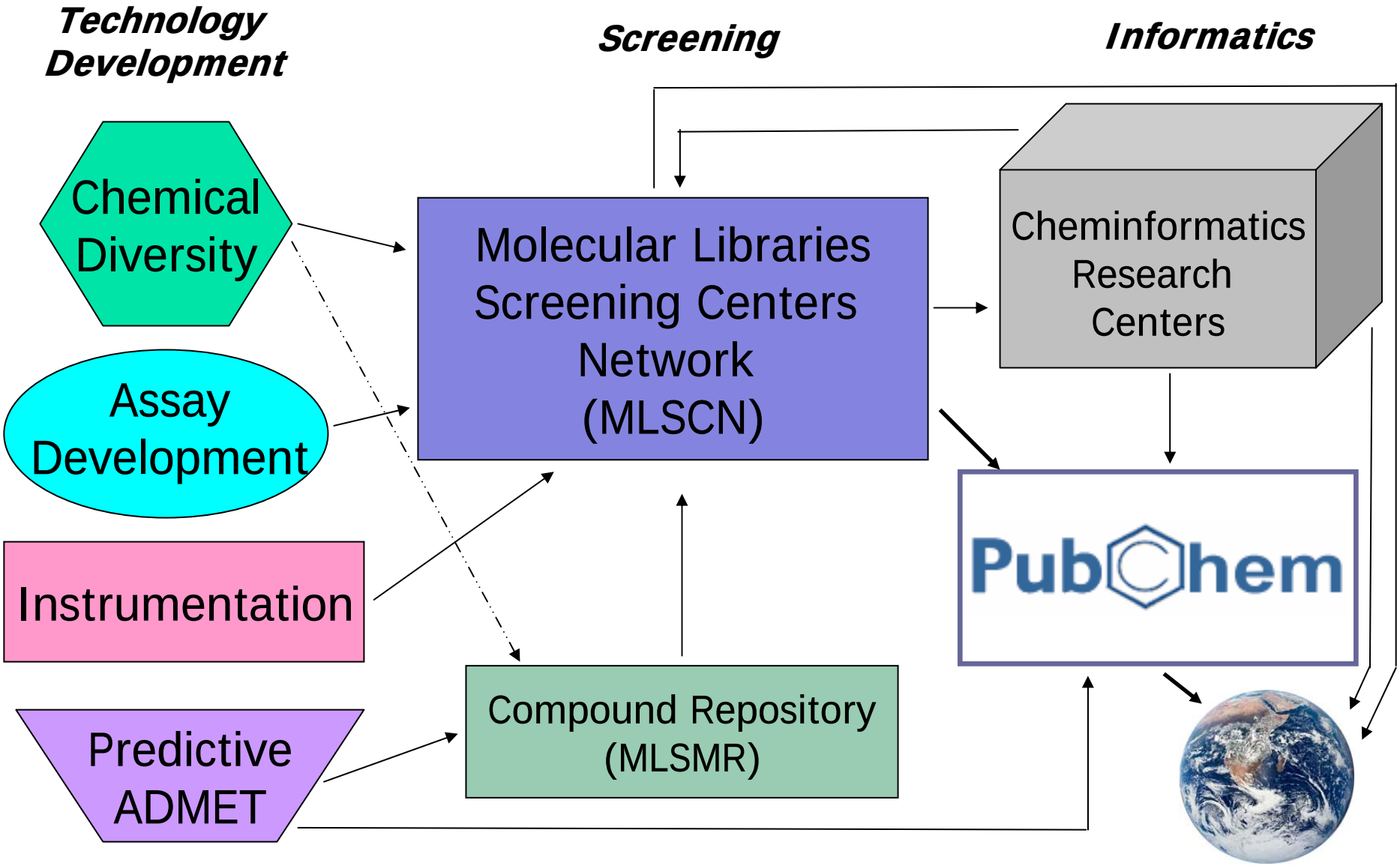
Molecular Libraries and Imaging Roadmap

- A trans-NIH initiative headed by:
  - Francis Collins, NHGRI
  - Tom Insel, NIMH
  - Rod Pettigrew, NIBIB
- Participating extramural staff members from 21 institutes and offices
  - CC
  - FIC
  - NCCAM
  - NCI
  - NCRR
  - NEI
  - NHGRI
  - NHLBI
  - NIA
  - NIAID
  - NIAAA
  - NIBIB
  - NICHD
  - NIDCD
  - NIDCR
  - NIDA
  - NIDDK
  - NIGMS
  - NIMH
  - NINDS
  - NLM

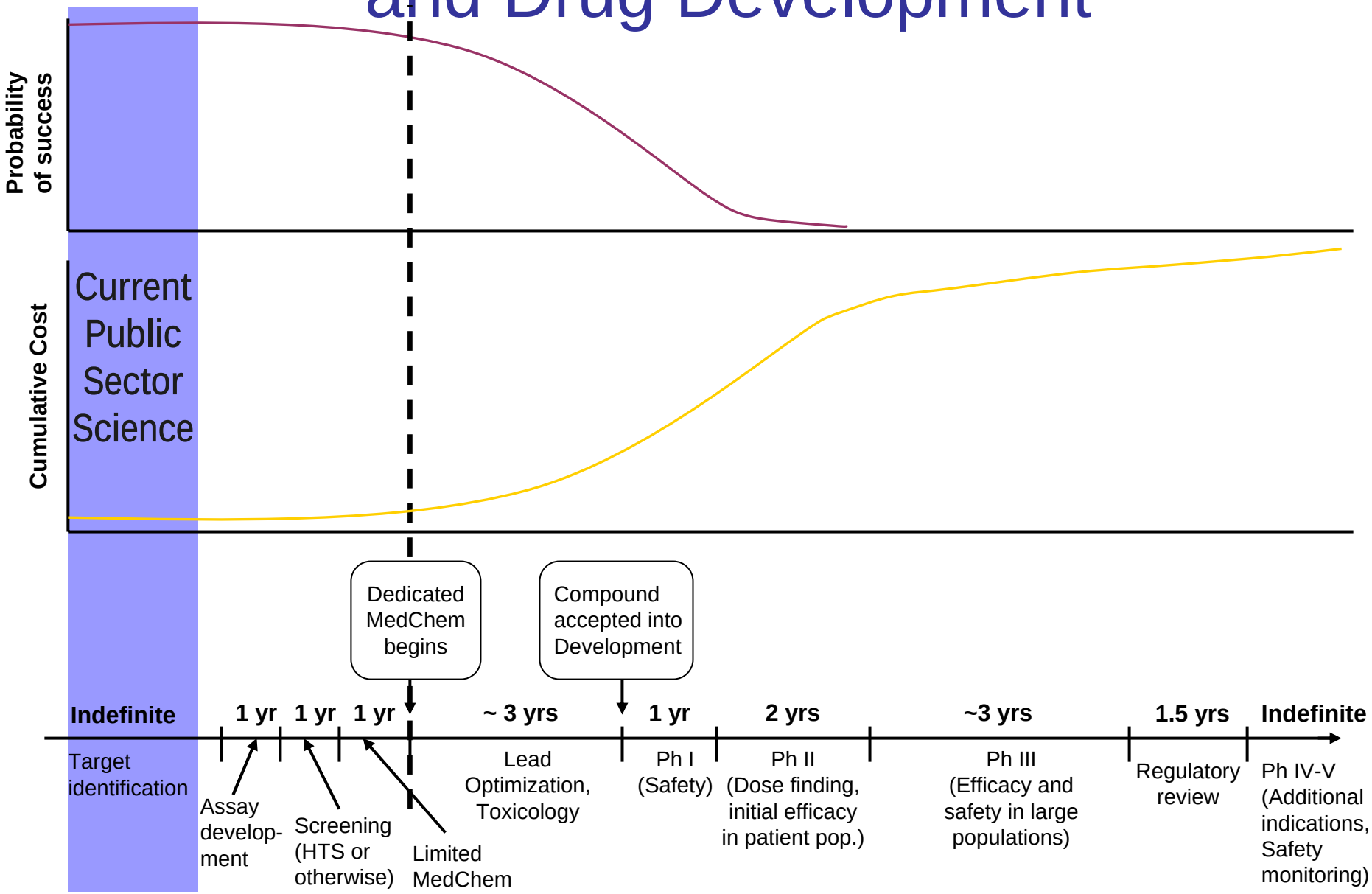
Molecular Libraries Roadmap Components

- Molecular Libraries Screening Center Network (MLSCN)
  - Small Molecule Repository (SMR)
  - NIH Chemical Genomics Center (NCGC, intramural)
  - Extramural Screening Centers
- Cheminformatics
  - PubChem
  - Cheminformatics Research Centers
- Technology Development
  - Chemical diversity
  - Assay development
  - HTS instrumentation
  - Predictive ADME/Toxicology

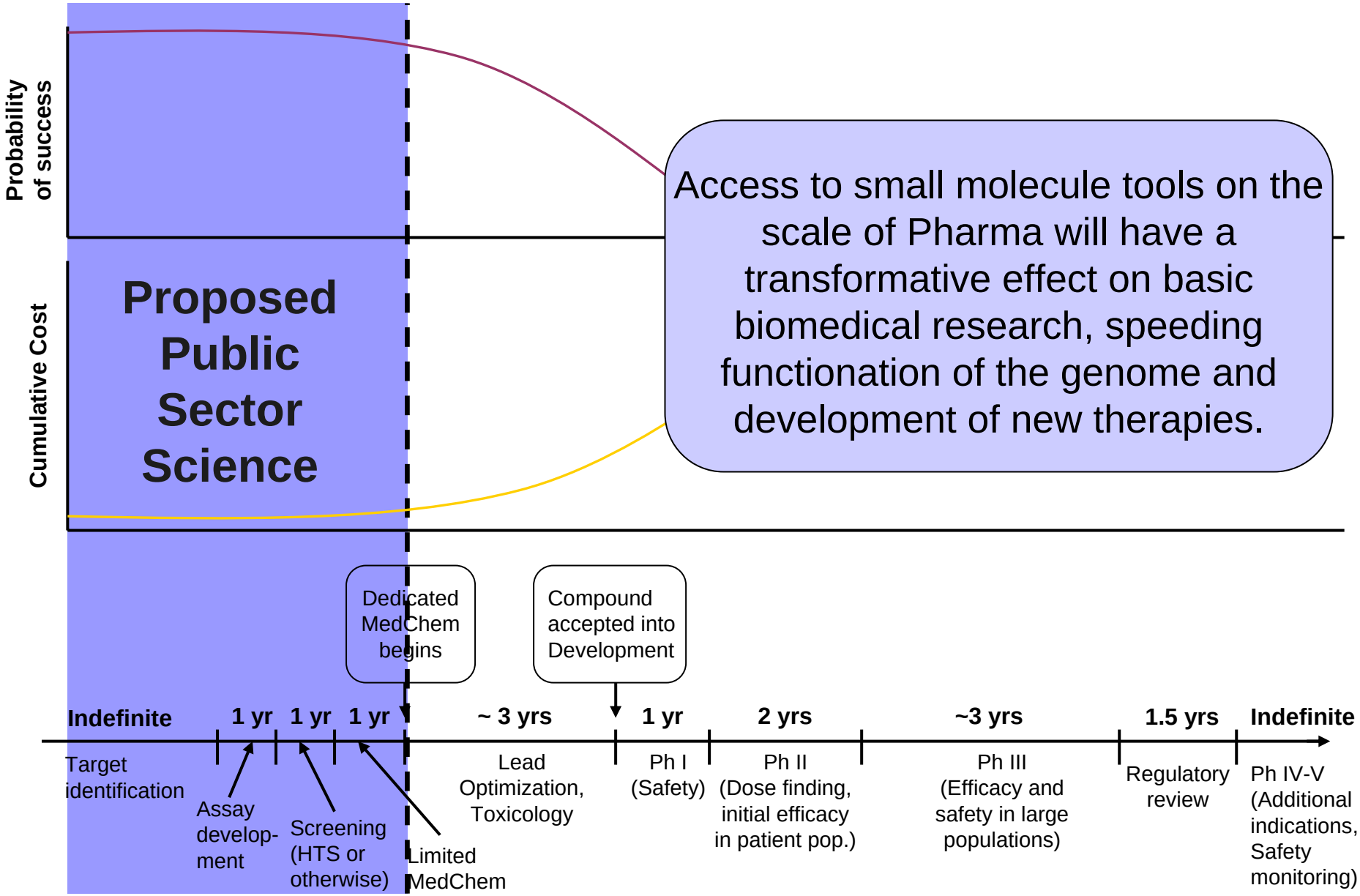
The Molecular Libraries Roadmap: An Integrated Initiative



The Interface of Chemical Genomics and Drug Development



The Drug Development Pipeline

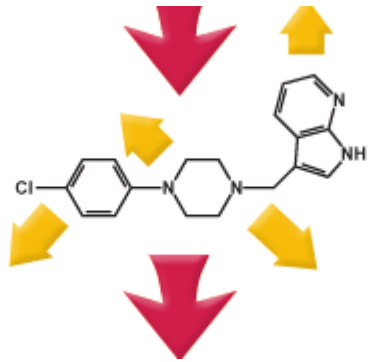


Molecular Libraries Screening Centers Network (MLSCN) is a Multidisciplinary Team Effort

HTS Assays from the Community

Compounds from the SMR

MLSCN Screening Centers Network



PubChem

NATIONAL INSTITUTES OF HEALTH

Goals of the MLSCN

- Establish a national HTS resource in the academic environment to improve the understanding of biology and disease mechanisms
  - Provide HTS approaches for identification of small organic molecules that are active in biological assays
  - Synthetic chemistry to improve the utility of small molecules as bioactive probes
  - Make HTS data publicly available in PubChem
  - Stimulate collaborations between biologists and chemists: assay providers, compound providers, and the MLSCN centers
    - Generate scientific publications, new research projects
  - Stimulate technology development
  - Provide outreach to the academic community

To Achieve the Goals of the MLSCN

- All data generated by the MLSCN will be deposited promptly upon data verification into PubChem
  - Data includes:
    - Assay descriptions/protocols,
    - Performance data for assays and compounds,
    - Primary data from HTS, data from secondary screening
    - Chemical structures/synthesis protocols for chemical analogs of hits and for probes
    - Biological activity of chemical analogs and probes

MLSCN RFA and Addendum: Guidance for Sharing of Data & Resources

- Community Resources
  - The usefulness of data and resources generated by the MLSCN will be of maximal benefit to public health if they are treated as a community resource and made publicly available
- Shared Resources include:
  - Data resulting from HTS of public domain compounds in the SMR
    - Compounds identified as “hits” are not likely to be immediately useful as research tools or as a final product
    - Screening data are deposited in PubChem with no delay
  - Synthetic chemistry and probe development
    - Generation of secondary libraries around a “hit”
      - Screening data is deposited into PubChem and compounds are made available in the Small Molecule Repository
  - Assay implementation
    - Assay descriptions and screening data are made available in PubChem
- Exceptions for immediate data release and sharing of resources will be considered
- http://grants.nih.gov/grants/guide/notice-files/NOT-RM-04-014.html

MLSCN Project Team Policy: Key Requirements for Data Sharing & IP

- Assays and improvements in assay methods may be patented
- Primary HTS by the MLSCN
  - Compounds from the SMR identified as hits from primary HTS are pre-competitive
  - Upon verification of hit activity for a compound, the screening results, compound structure, and performance data on the hit must be promptly deposited in PubChem
  - MLSCN centers and assay providers will define criteria for data verification for an assay and receive acceptance from the Scientific Program Managers prior to commencing work on the assay
- http://www.nimh.nih.gov/dnbbs/datasharing-ip.pdf

MLSCN Project Team Policy: Key Requirements for Data Sharing & IP

- Secondary screening
  - Results from secondary screening of initial hits and chemical analogs of hits must be promptly deposited in PubChem
- Synthetic chemistry and chemical probe development
  - Development of chemical analogs of hits should be made solely in order to meet the specifications of a chemical probe
  - Once the specifications of a chemical probe are achieved, further optimization ceases within the MLSCN
    - Definition of a chemical probe
  - Primary and secondary screening results, and chemical synthesis protocols and data, on all rounds of chemical optimization of hits into probes must be promptly deposited in PubChem
  - Following deposition in PubChem of the chemical probe data, the centers are free to pursue chemical optimization independently of the MLSCN program and MLSCN funds

MLSCN Project Team Policy: Exceptions for Two Unique Circumstances

- If prompt disclosure of screening data is determined to provide an unfair advantage to scientific competitors of the assay provider
  - A waiver of the requirement of prompt deposition of screening data will be considered by the NIH Project Team to allow early publication of the screening results
  - If approved, the delay in data deposition would not exceed 60 days
- If a compound is determined to be a drug candidate without further chemical optimization
  - A waiver of the requirement of prompt deposition of screening data will be considered by the NIH Project Team to allow filing of a patent
  - If approved, the delay in data deposition would not exceed 60 days

Molecular Libraries Screenings Centers Network (MLSCN)

PI Name	Institution Name	Title
AUSTIN, CHRIS	NIH	The NIH Chemical Genomics Center
DIAMOND, SCOTT	UNIVERSITY OF PENNSYLVANIA	The Penn Center for Molecular Discovery
DINGLEDINE, RAYMOND	EMORY UNIVERSITY	Emory Chemistry-Biology Center in the MLSCN
LAZO, JOHN	UNIVERSITY OF PITTSBURGH AT PITTSBURGH	University of Pittsburgh Molecular Libraries Screening Center
PIAZZA, GARY	SOUTHERN RESEARCH INSTITUTE	Southern Research Molecular Libraries Screening Center
REED, JOHN	THE BURNHAM INSTITUTE	San Diego Center for Chemical Genomics
ROSEN, HUGH	THE SCRIPPS RESEARCH INSTITUTE	Scripps Research Institute Molecular Screening Center
ROTHMAN, JAMES	COLUMBIA UNIVERSITY MEDICAL CENTER	MLSCN Center at Columbia University
SKLAR, LARRY	UNIVERSITY OF NEW MEXICO ALBUQUERQUE	New Mexico Molecular Libraries Screening Center
WEAVER, C. DAVID	VANDERBILT UNIVERSITY	Vanderbilt Screening Center for GPCRs, Ion Channels, and Transporters

MLSCN Nationwide Network

MLSCN Center
at Columbia University

NIH Chemical Genomics
Center (NCGC)

University of Pittsburgh Molecular
Libraries Screening Center

The Penn Center for
Molecular Discovery

Scripps Florida
Molecular Screening Center

Emory Chemistry-Biology
Center in the MLSCN

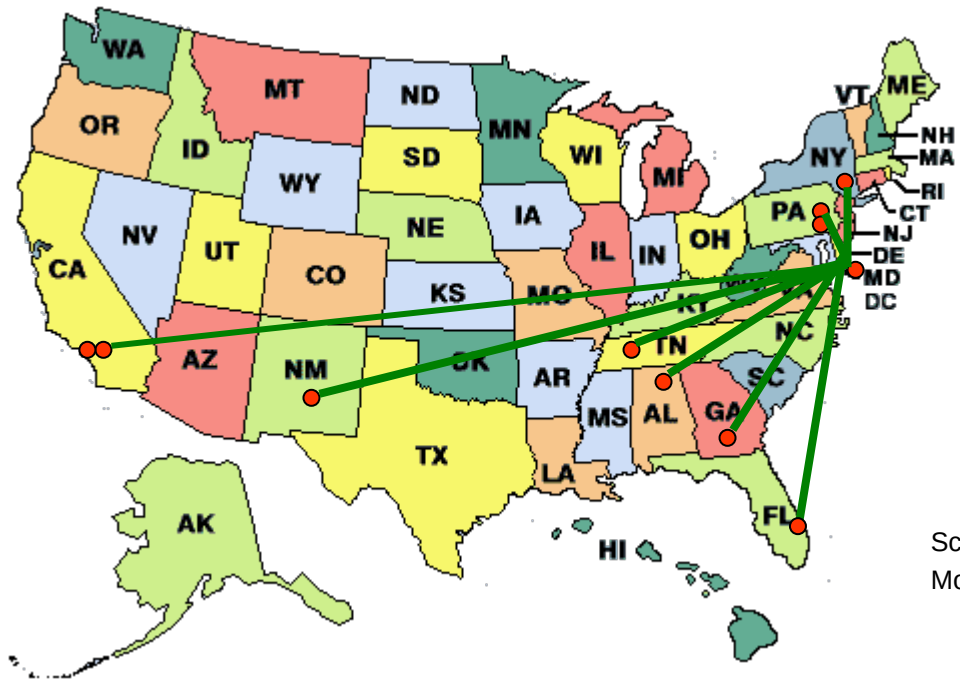
Southern Research Molecular Libraries
Screening Center (SRMLSC)

Vanderbilt Screening Center for
GPCRs, Ion Channels, and Transporters

New Mexico Molecular Libraries
Screening Center

San Diego Center for
Chemical Genomics

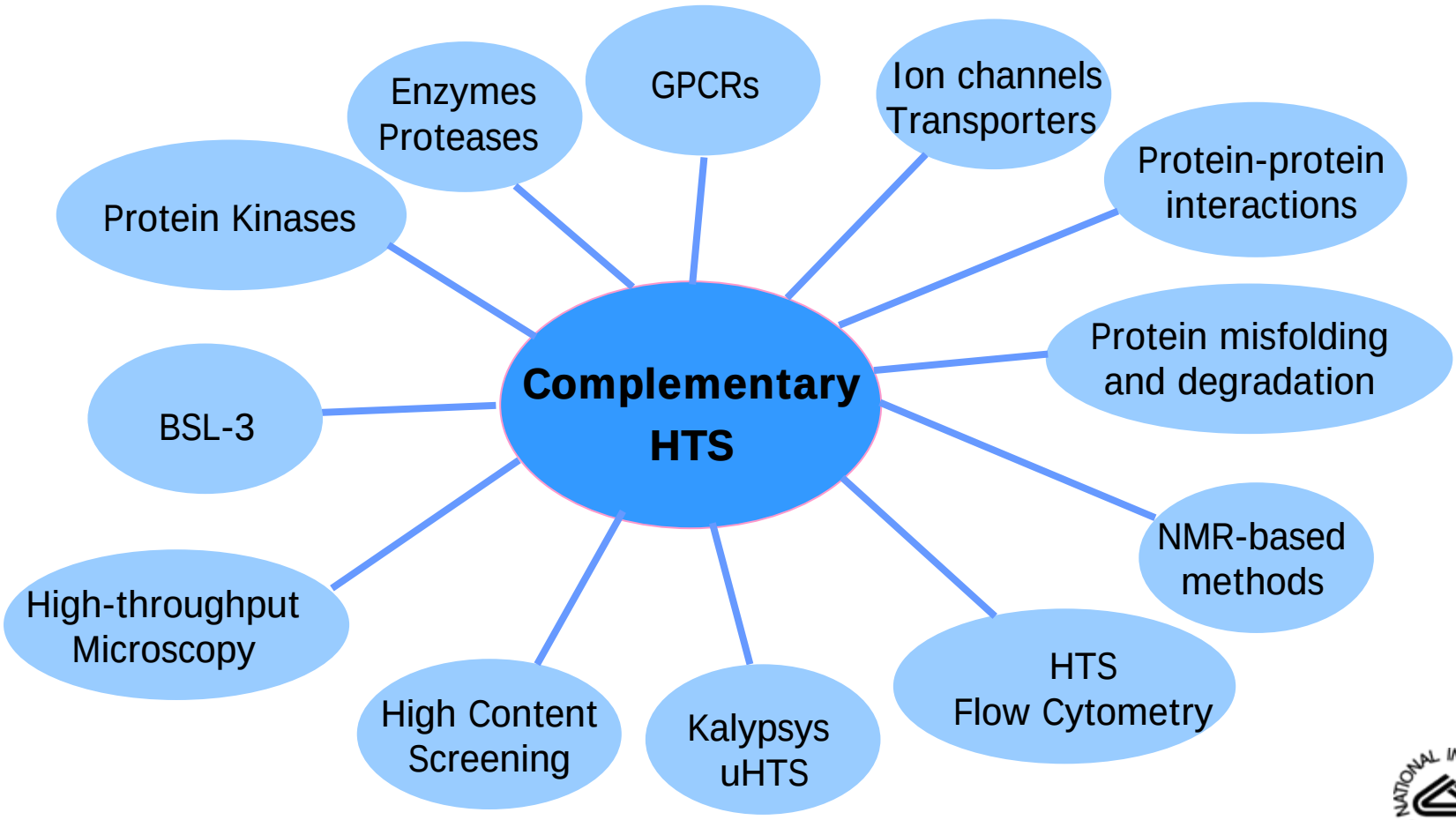
Scripps Research Institute
Molecular Screening Center



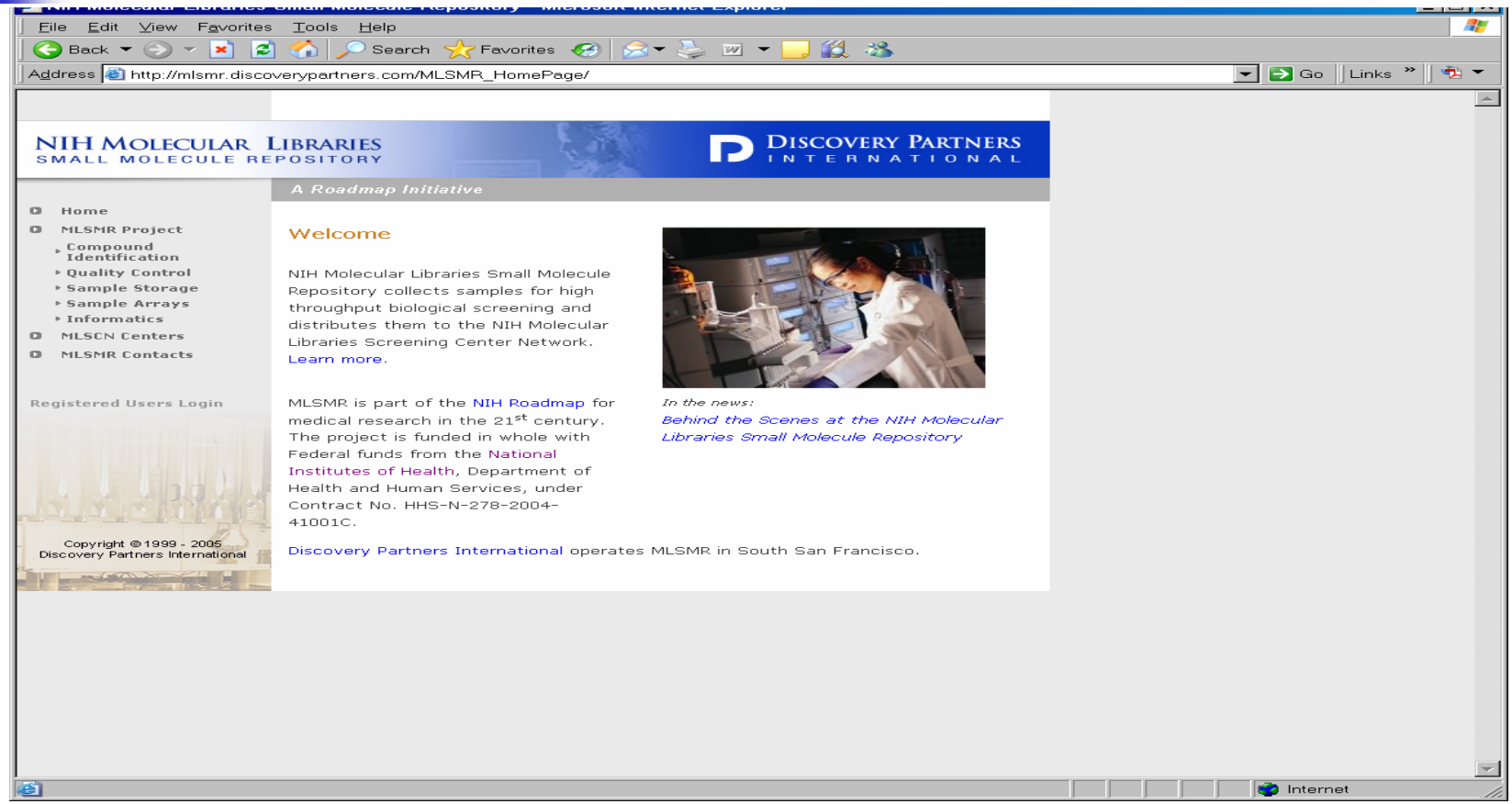
Overall HTS Technology of the MLSCN

- Diverse screening platform technologies
  - Cell based imaging, flow cytometry, microarray screening, ultraHTS, protein-protein, whole organism
- Detection systems
  - Luminescence, fluorescence, FLIPR, FRET, SPA, absorbance, ELISA
  - Microscopy based imaging, spectrophotometric
  - Microarray, RT-PCR

Overall Capabilities of the MLSCN



The NIH Molecular Libraries Small Molecule Repository (MLSMR)



http://mlsmr.discoverypartners.com/MLSMR_HomePage/

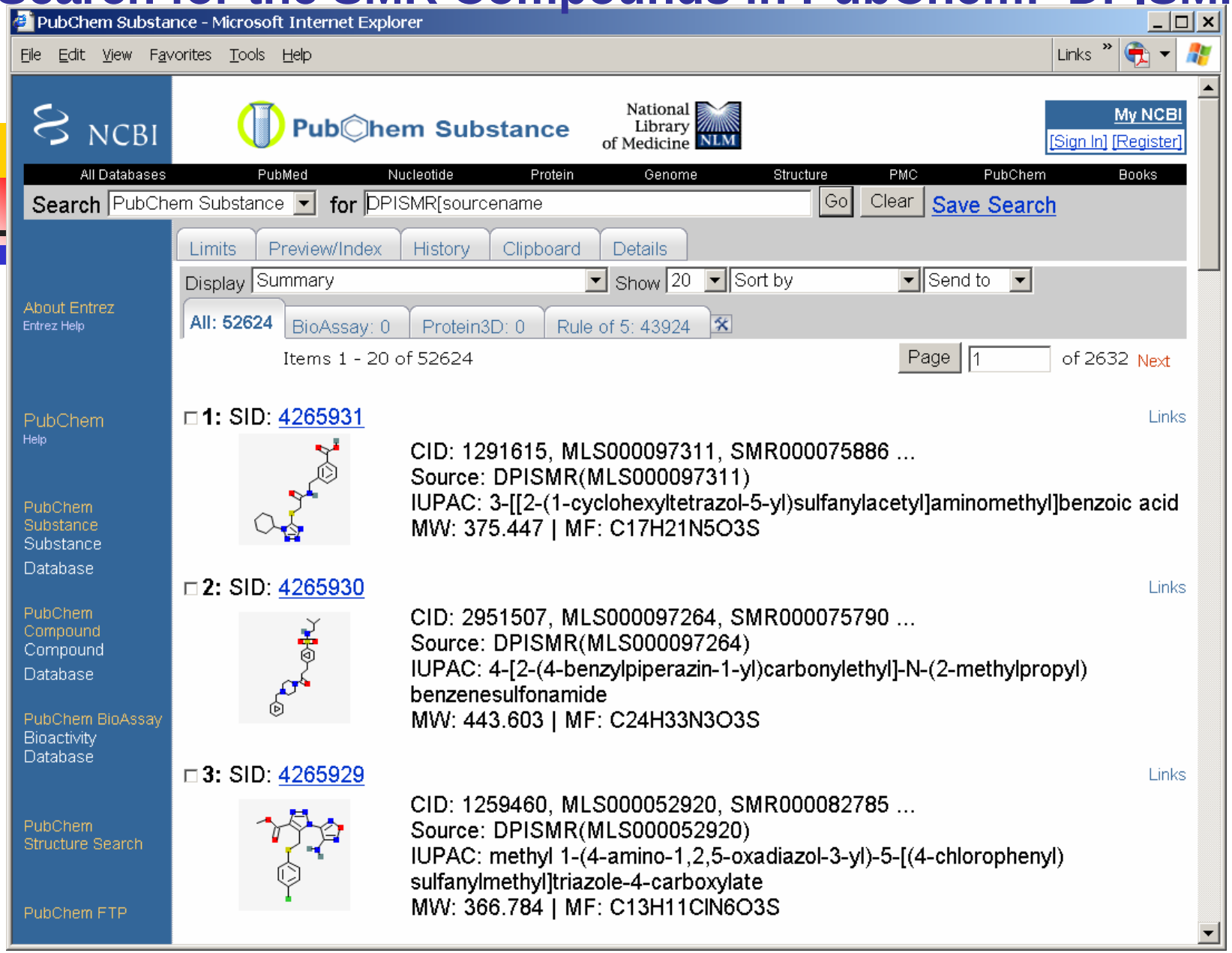
Building the MLSMR

- Initial set of 80,000 compounds
  - Compound Selection Criteria
    - Purity > 90%
    - Stock of > 10 mg
    - Lipinski Rule of 5, if applicable
    - Solubility of 20 ug/ml based on AlogS
    - No undesirable functional groups

Molecular Libraries Small Molecule Repository (MLSMR)

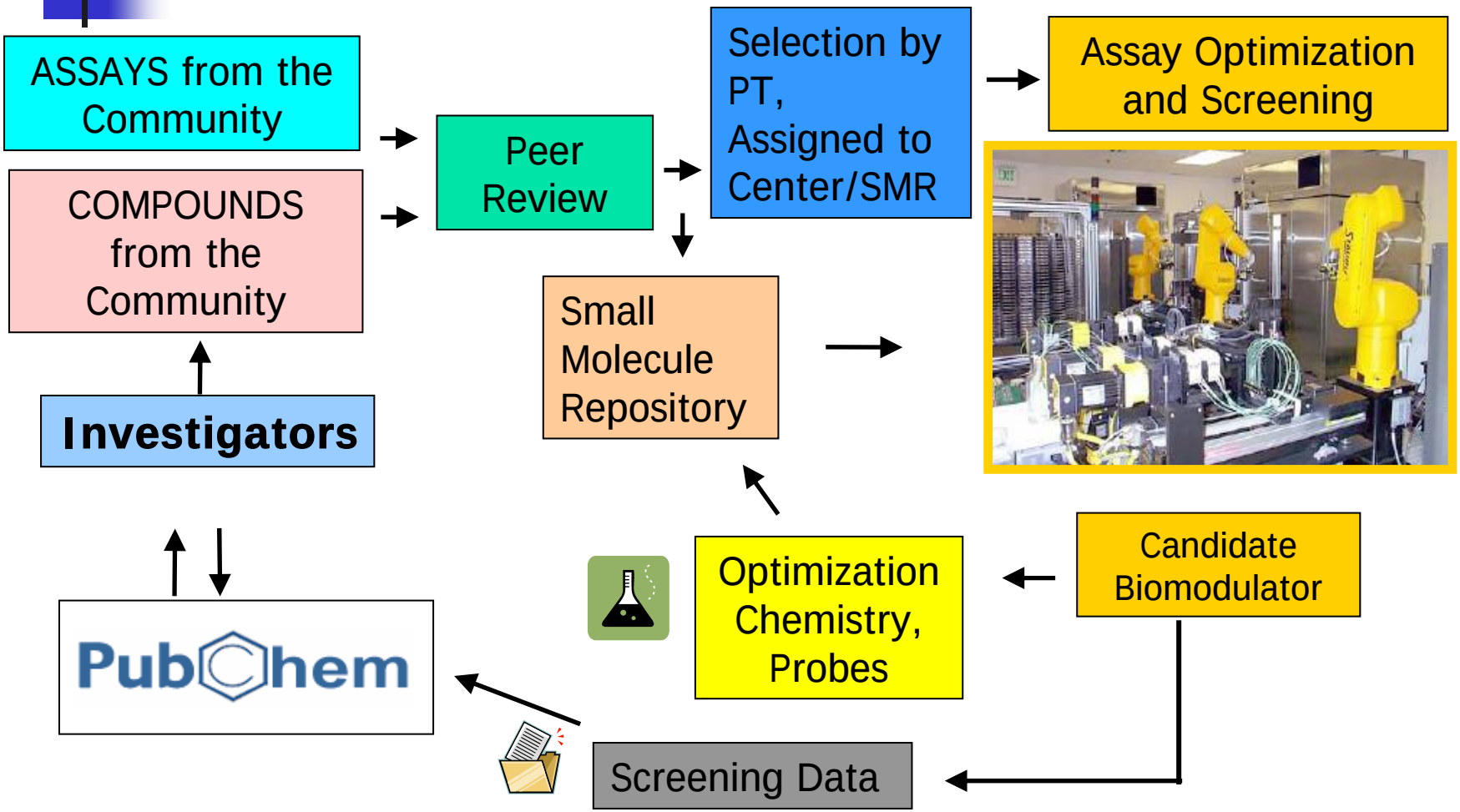
- Initial set of 80,000 compounds purchased from commercial vendors
  - Targeted Libraries
    - Active ingredients of FDA approved drugs
  - Diverse compounds
  - Natural products
- Expanding the compound collection
  - Solicitation of compounds from academia, biotech companies, and pharma (in process)
  - NIGMS Chemical Methodologies Library Development (CMLD) centers
  - Molecular Libraries Chemical Diversity initiatives
  - Strategy being developed for acquisition of second 100,000 compounds

To Search for the SMR Compounds in PubChem: DPISMR

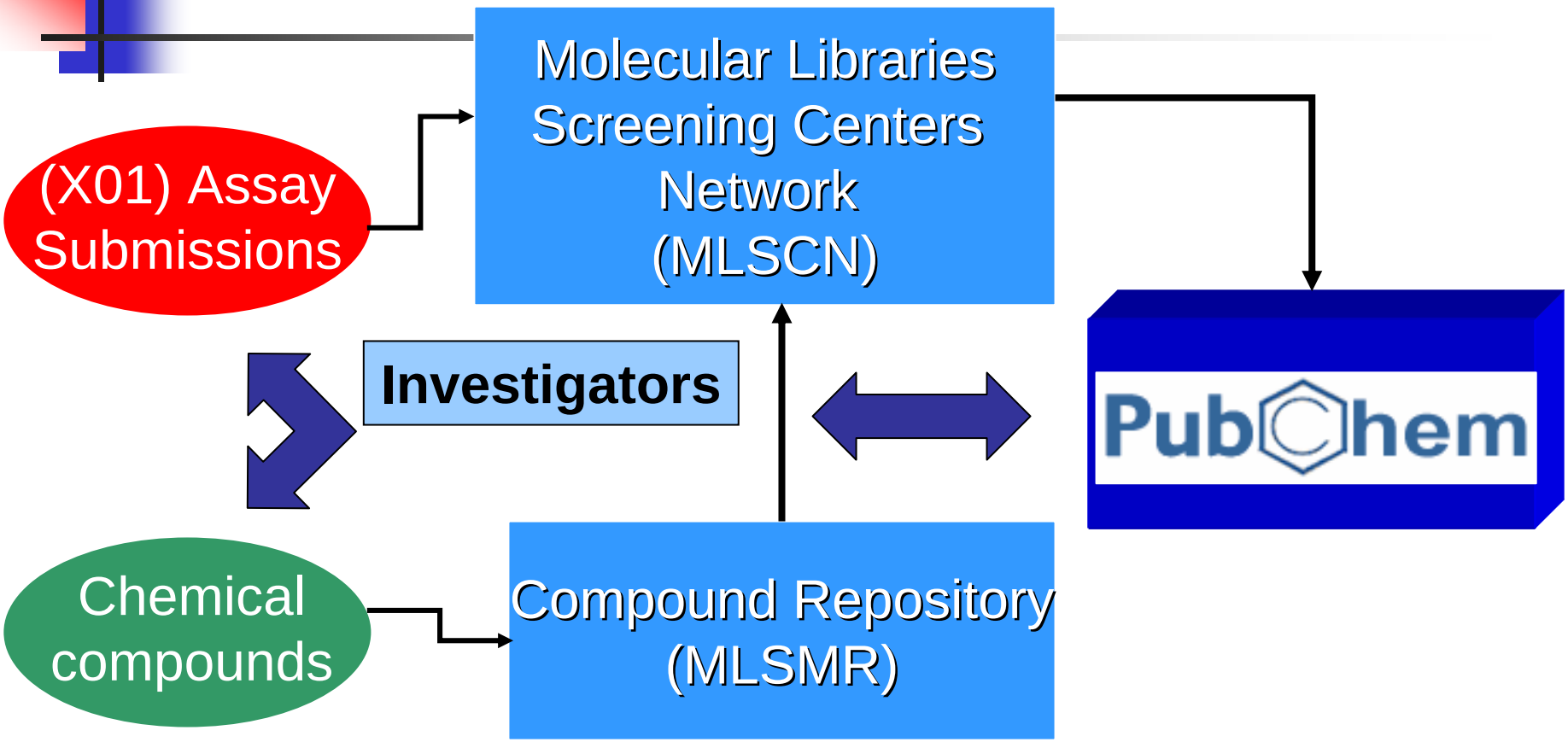


http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pcsubstance&term=DPISMR[sourcename

MLSCN Resources



Opportunities for Accessing MLSCN Resources



Incentives for Investigators to Interact with the MLSCN

- Access to the resources of the MLSCN Network of Screening Centers through the X01 mechanism
  - Diverse screening platform technologies
    - Cell based imaging, flow cytometry, microarray screening, ultraHTS, protein-protein, whole organism
  - Detection systems
    - Luminescence, fluorescence, FLIPR, FRET, SPA, absorbance, ELISA
    - Microscopy based imaging, spectrophotometric
    - Microarray, RT-PCR
- Availability of research tools/probes and biological data deposited in PubChem

Incentives for Investigators to Interact with the MLSCN

- Access to compounds in the Small Molecule Repository for HTS through the MLSCN
  - Current library of 80,000 compounds purchased from commercial vendors
    - Targeted Libraries
      - Active ingredients of FDA approved drugs
    - Diverse compounds
    - Natural products
  - Ongoing expansion of chemical diversity of the repository

    Solicitation of compounds from academia, biotech companies, and pharma (in process)
    - NIGMS Chemical Methodologies Library Development (CMLD) centers
    - Molecular Libraries Chemical Diversity initiatives
    - Acquisition of second 100,000 compounds

Opportunities and Incentives for Investigators to Interact with the MLSCN

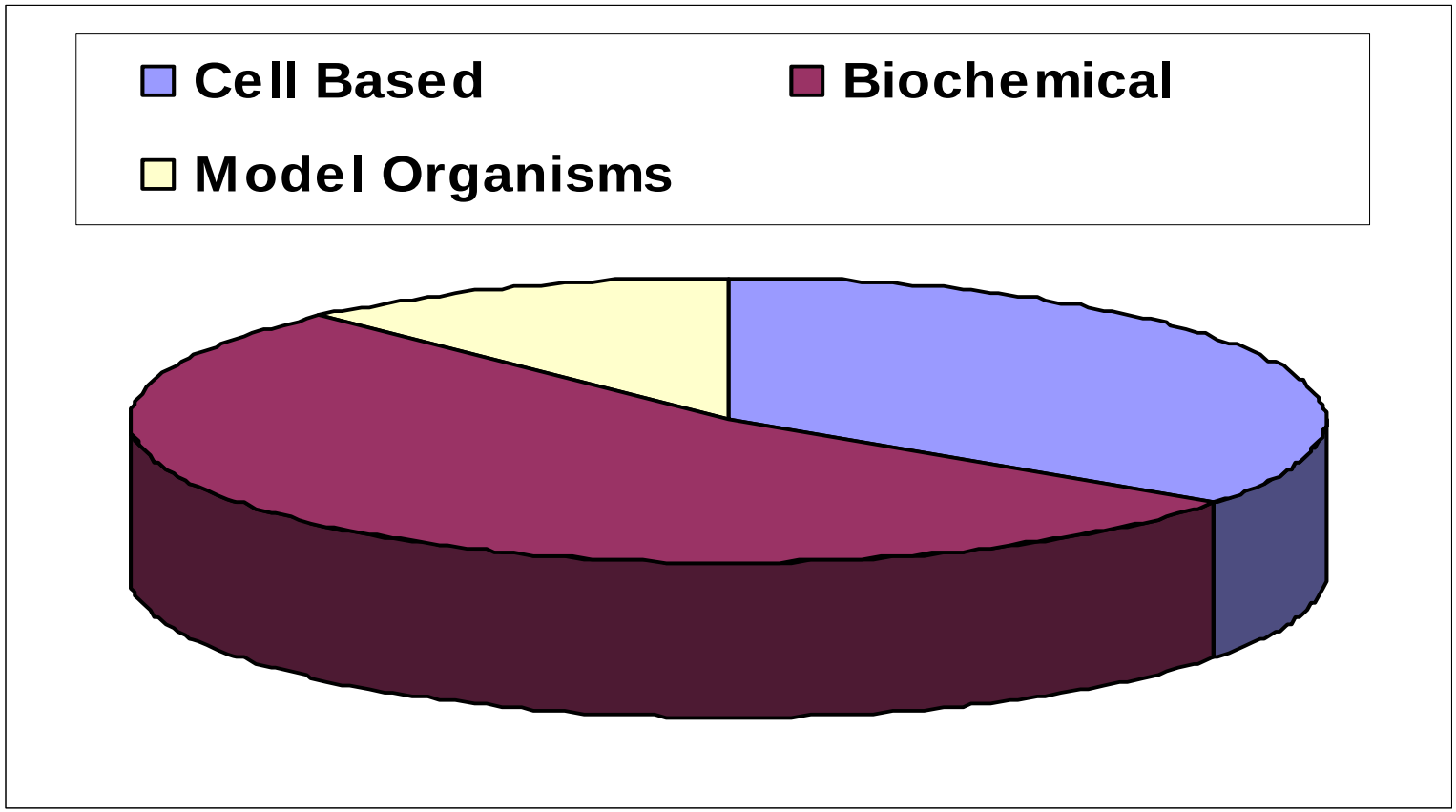
- Process for access to the resources of the network of the ten MLSCN centers
  - Submission of X01 assay applications; if selected for implementation:
    - Assay adapted for HTS
    - Screened against SMR
    - Collaboration with assay provider for secondary screening of confirmed hits
    - Chemical optimization of confirmed hits
  - Submission of compound applications (to be implemented) for entry into the Repository
    - Compounds screened in HTS assays in the MLSCN

MLSCN HTS Assay Solicitation: X01 Resource Access Award

- The MLSCN accepts assays for HTS from the broad research community, including
  - Public or private institutions and agencies of the Federal government
- The MLSCN intends to select approximately 100-200 assays per year for implementation within the network of screening centers. It is expected that each screening center will conduct 10-20 HTS assays annually
- Applications are invited from investigators who have developed innovative assays for use both in basic research and in therapeutics development programs

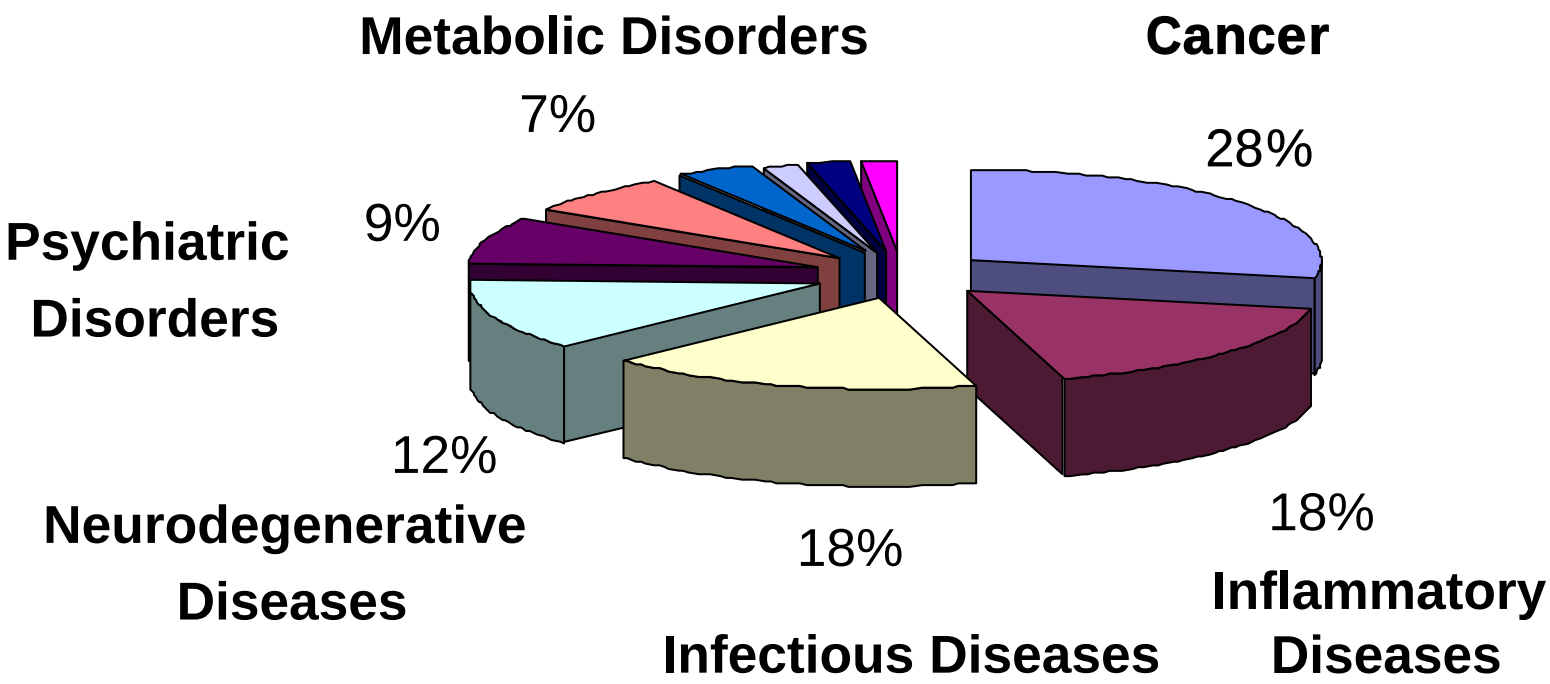
MLSCN HTS Assays to Date

Assay Formats



MLSCN HTS Assays to Date

Disease Relevance



Opportunities for Investigators to Interact with the MLSCN

- Access to and mining of data deposited in PubChem
  - Primary and secondary biological screening data
  - Assay descriptions and protocols
  - Synthesis protocols for optimization chemistry and research tools or probes generated by the MLSCN
  - Chemical structures of compounds in the Small Molecule Repository in addition to many other compound entries

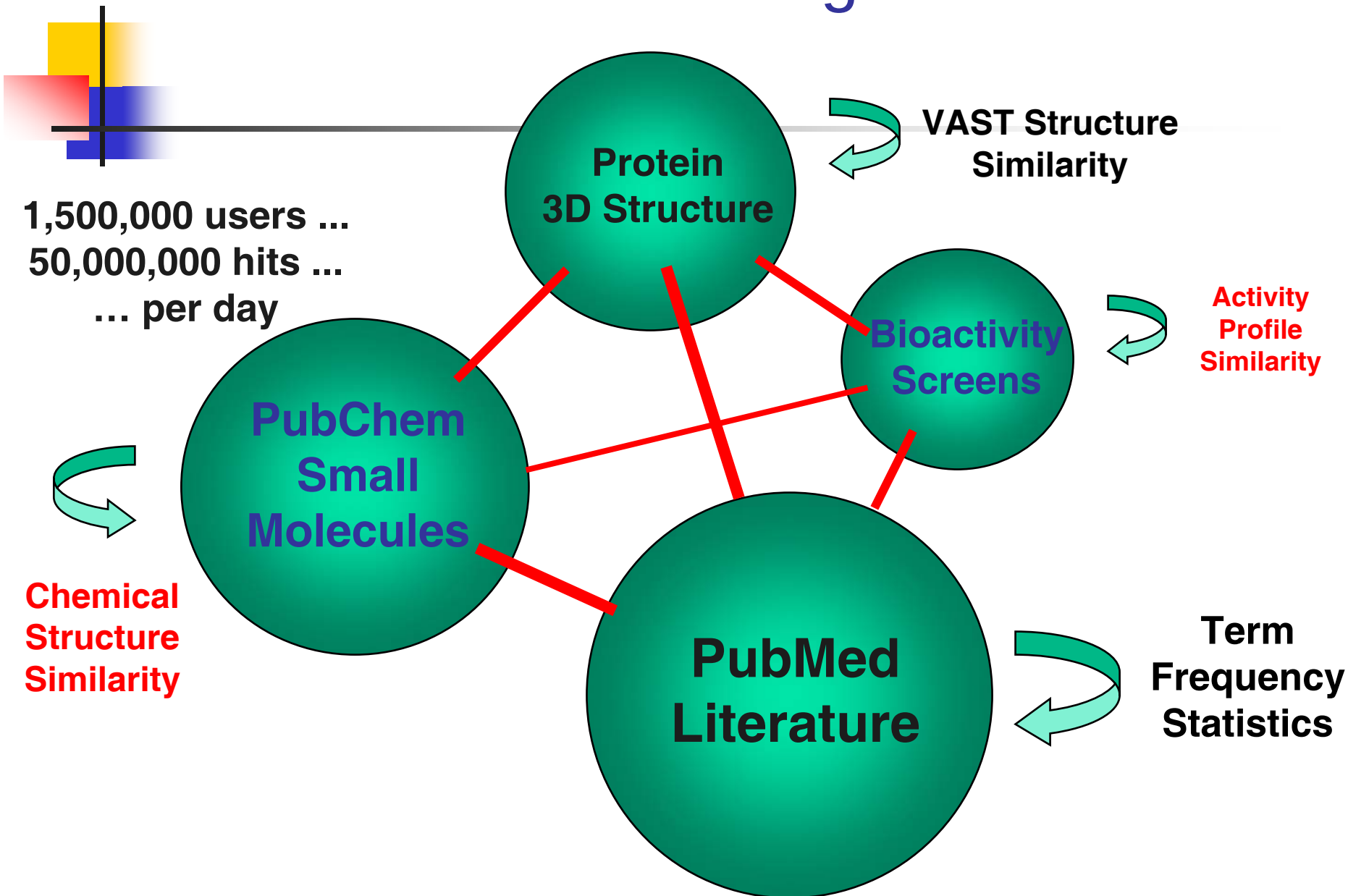
Molecular Libraries: PubChem

- Public sector chemical database developed by the NIH National Center for Biomedical Informatics
  - Fully linked to other NCBI Entrez databases of genes, proteins, Medline
  - Coordinates bioassay data deposition from the MLSCN
  - Provides support to investigators for bioassay data deposition
  - Went live in September 2004
    - http://pubchem.ncbi.nlm.nih.gov/

PubChem

- Entrez links and neighbors:
  - PubMed-PubChem linkage
  - Unique chemical structures
    - PubChem structure neighbors
    - Structure sketching for structure searches
  - PubChem bioassay links
- Contents as of March 2006:
  - Approximately 7,848,000 substance records
  - Deposited by > 25 government, academic, and commercial organizations
  - Approximately 5,270,000 unique chemical structures
  - Approximately 200 bioassay data sets deposited
    - > 3,158,000 bioassay test results

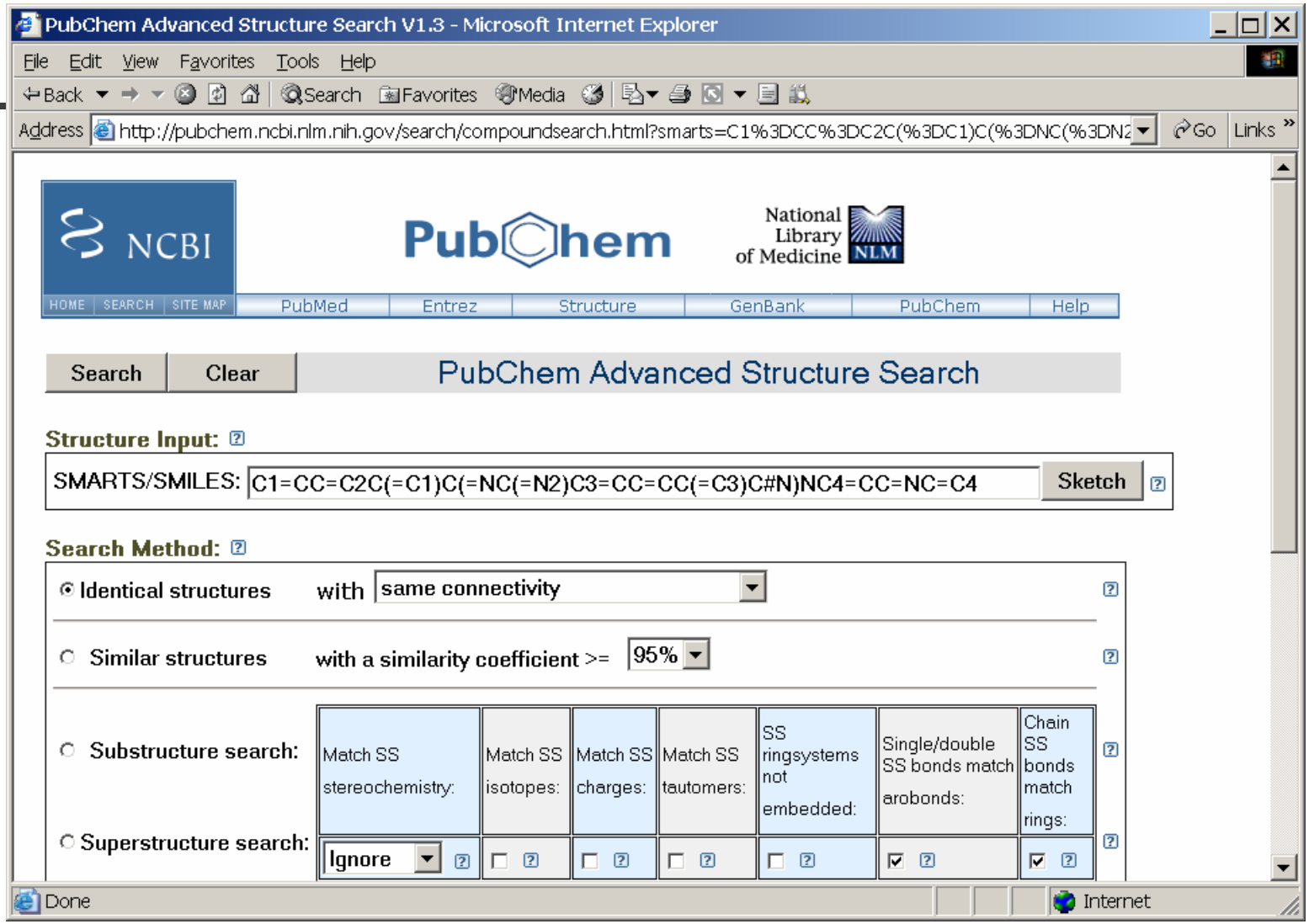
PubChem
Entrez Links and Neighbors



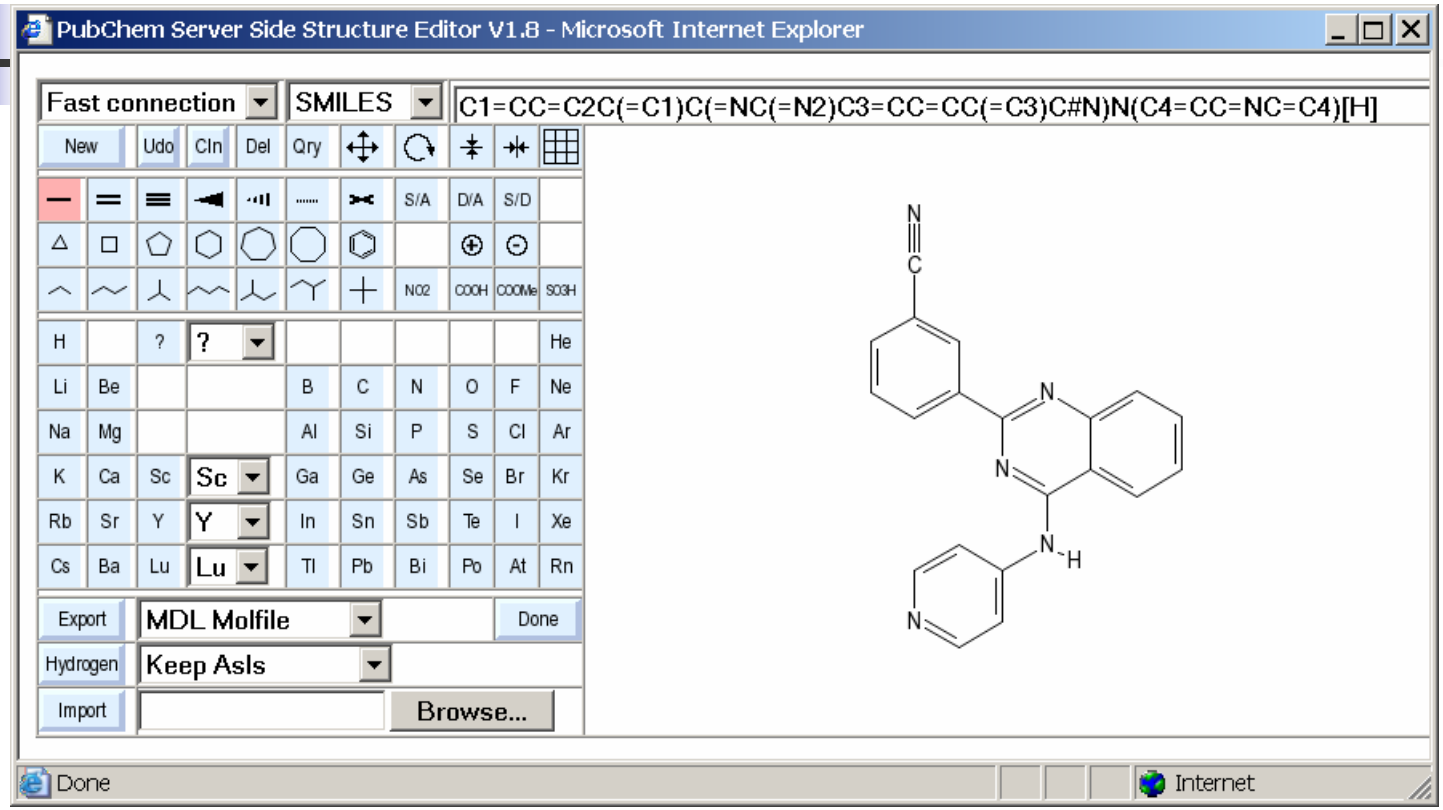
PubChem Information Sources

- Contributed substance records
- Contributed biological screening results
- Contributed links to other Entrez databases
- Links created by PubMed indexing
- Computed similarities between records

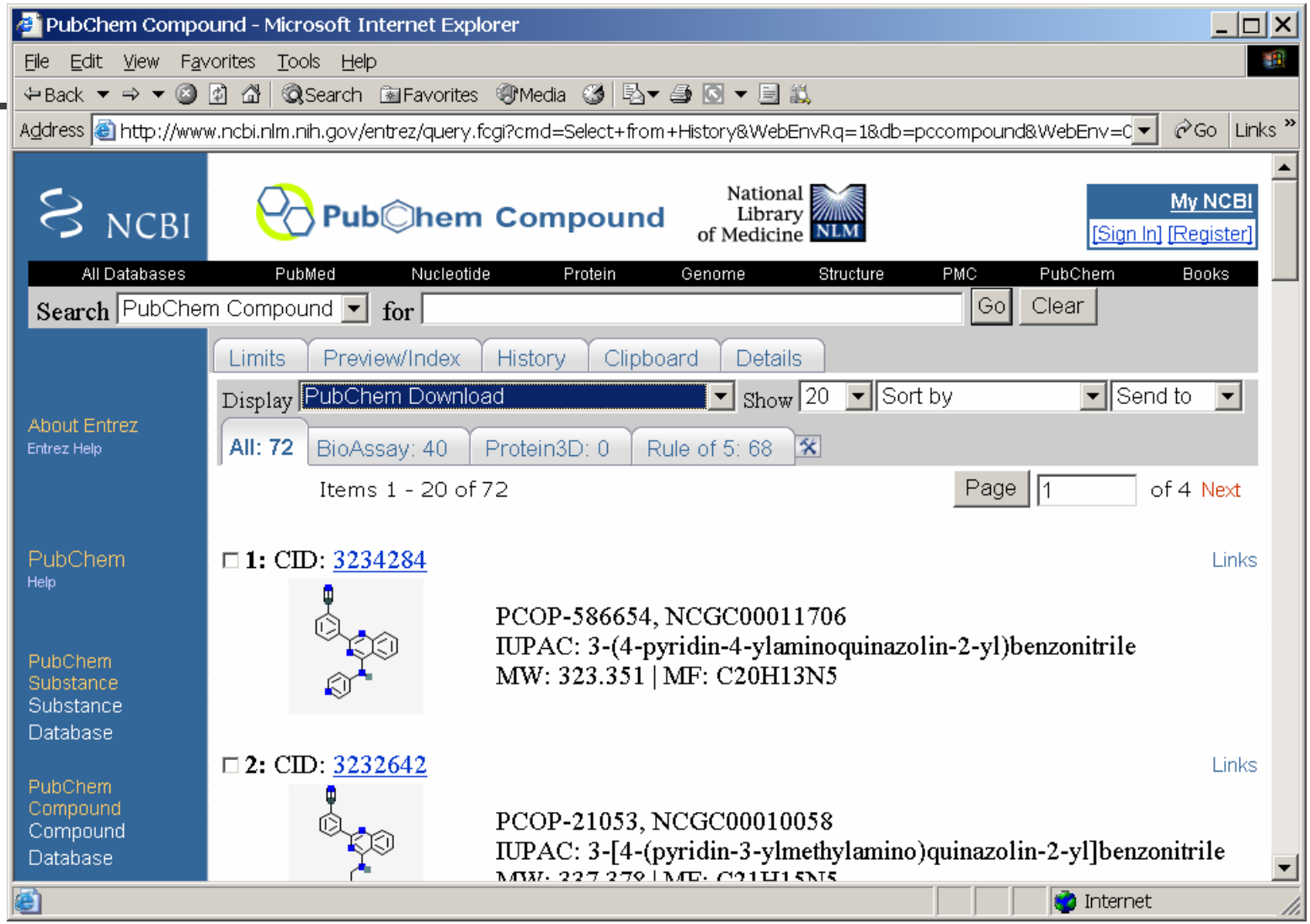
PubChem Structure Search Tool



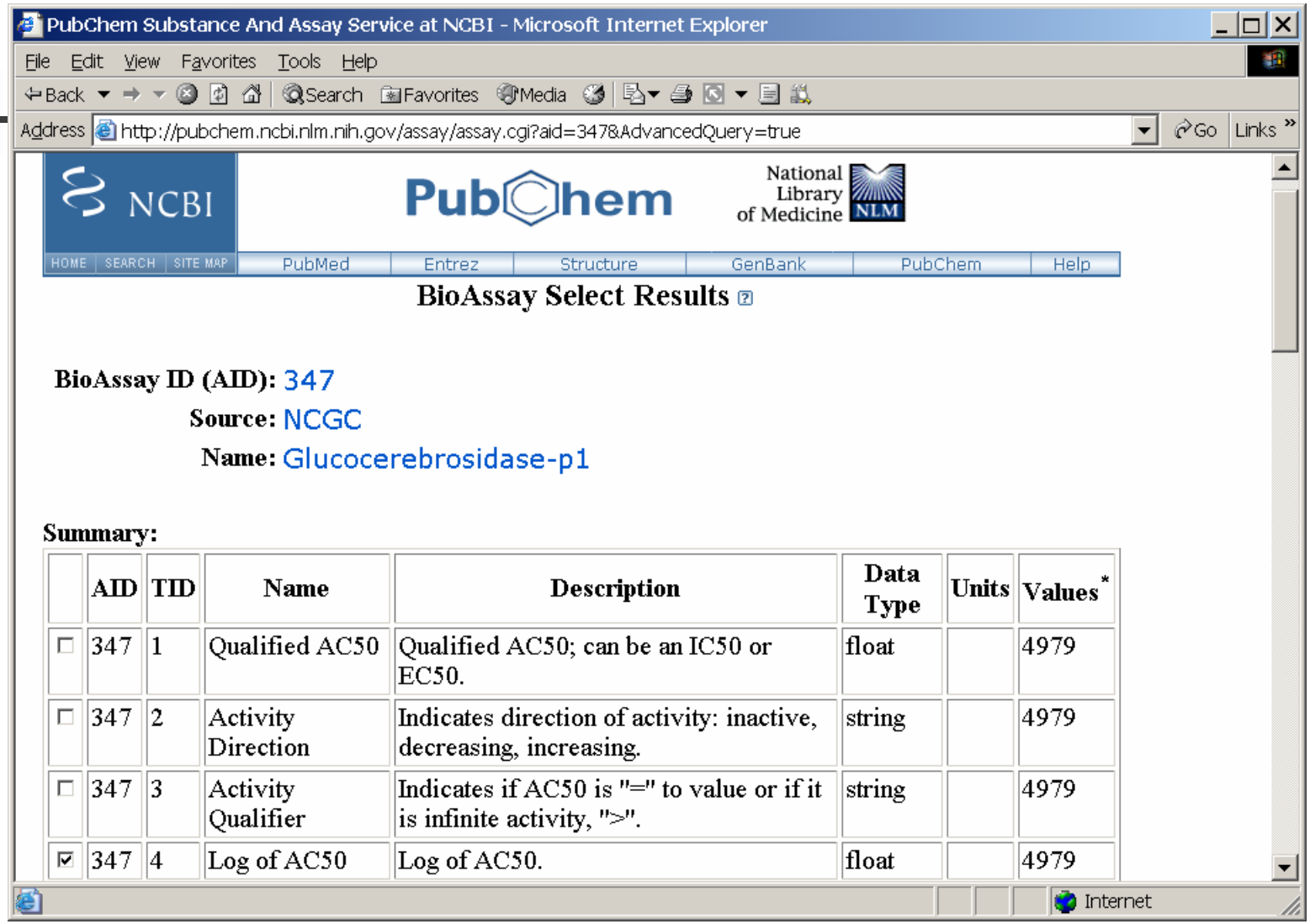
PubChem Sketcher and Other Options



PubChem Search Results
Pass to Entrez

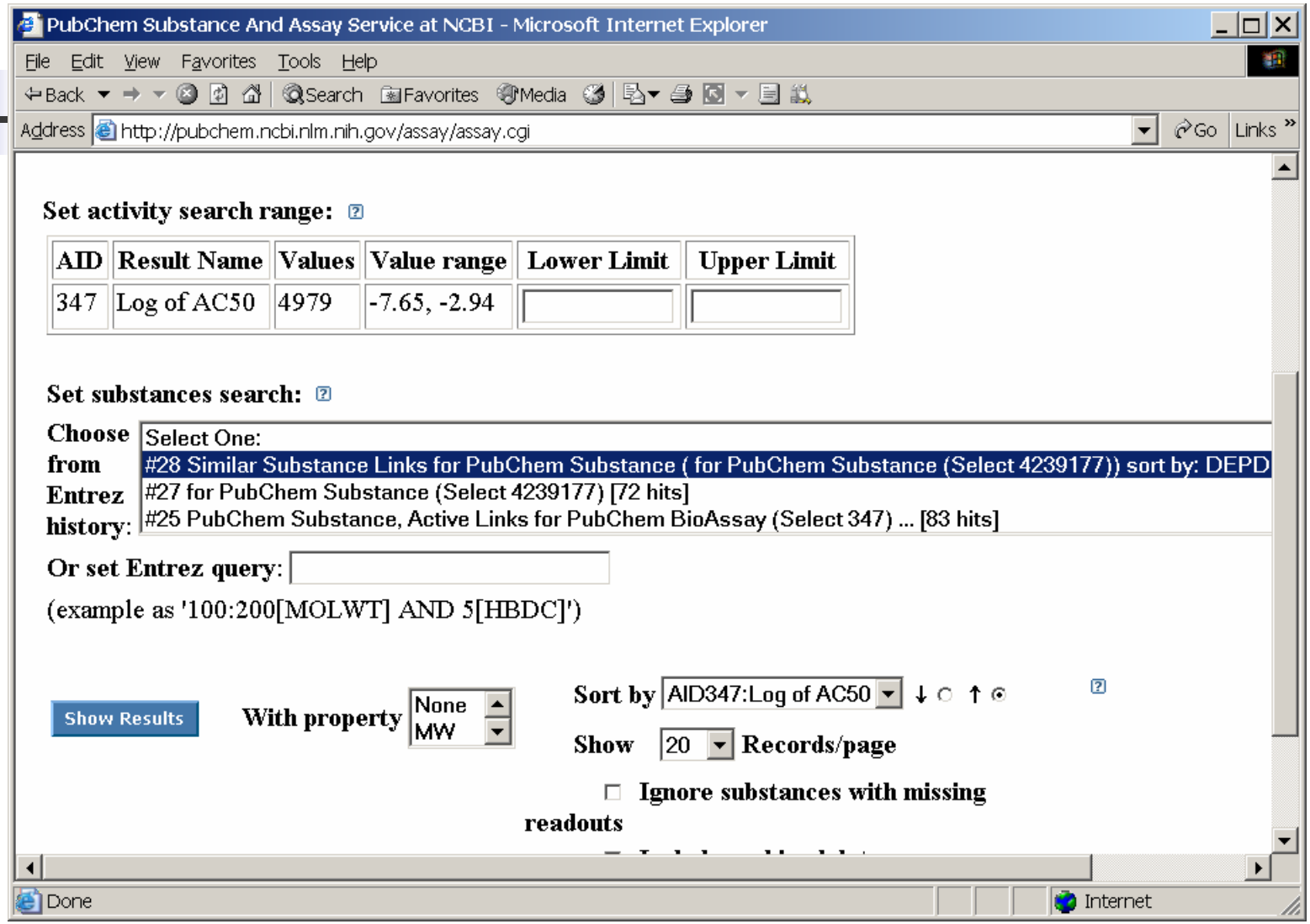


Select PubChem BioAssay Results



	AID	TID	Name	Description	Data Type	Units	Values*
☐	347	1	Qualified AC50	Qualified AC50; can be an IC50 or EC50.	float		4979
☐	347	2	Activity Direction	Indicates direction of activity: inactive, decreasing, increasing.	string		4979
☐	347	3	Activity Qualifier	Indicates if AC50 is "=" to value or if it is infinite activity, ">".	string		4979
☑	347	4	Log of AC50	Log of AC50.	float		4979

For Selected PubChem Substances



Molecular Libraries
Technology Development Initiatives

- Cheminformatics Research Centers
- Chemical diversity
- Assay development
- HTS instrumentation
- Predictive ADME/Toxicology

Cheminformatics Research Centers
Technology Development

- Substantial unmet need for publicly available cheminformatics research tools
  - Virtual screening, virtual synthesis, other applications, R&D on new tools
- Exploratory Centers for Cheminformatics Research, RFA RM-05-012
  - FY05: 6 Exploratory Centers awarded
    - P20s, 2 years

Chemical Diversity
Technology Development Supporting the MLSCN and SMR

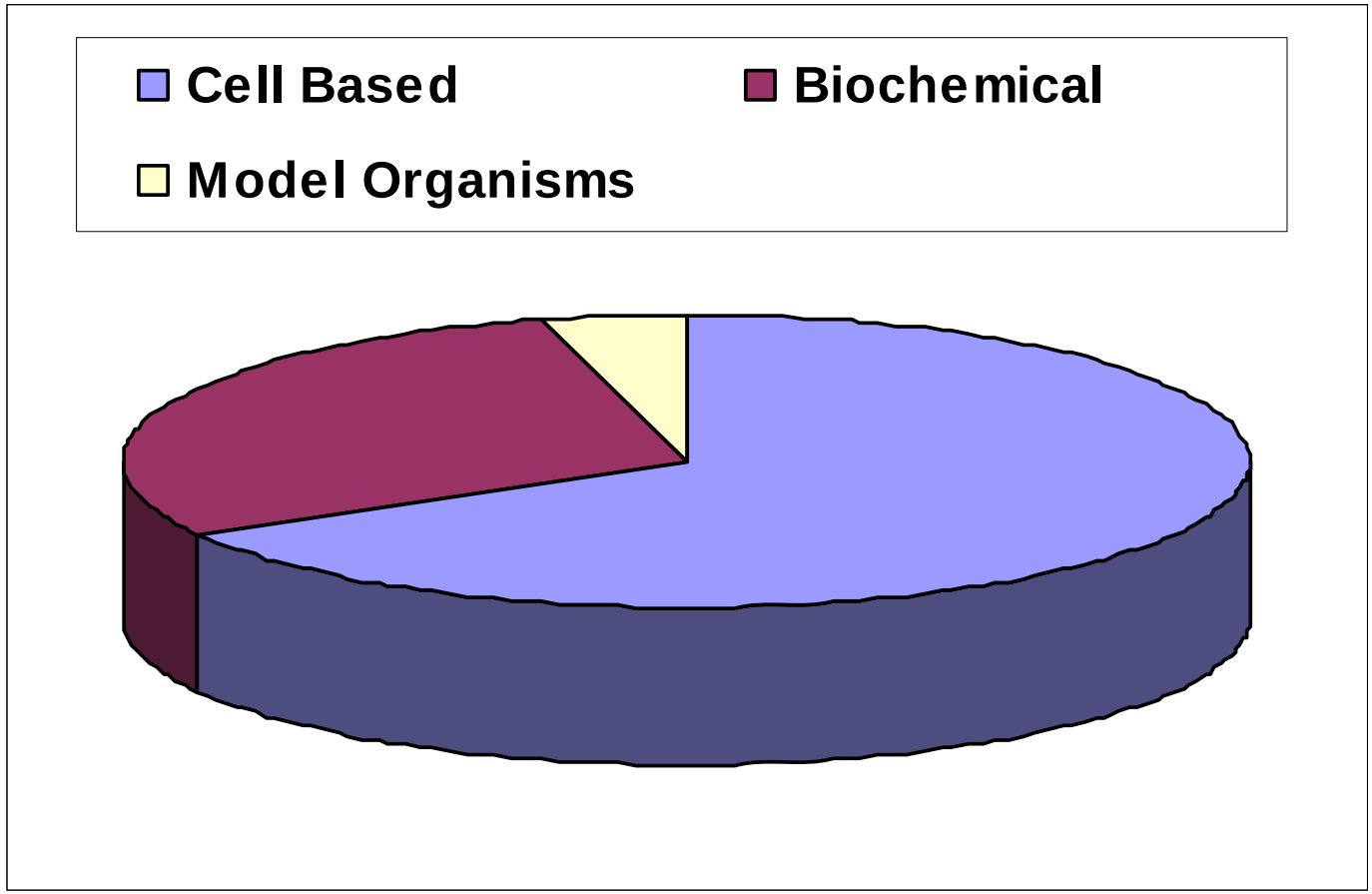
- Pilot-Scale Libraries for High-Throughput Screening (RFA-RM-05-014)
  - Rationale: Probes for novel proteins require novel chemical structures
  - Pilot chemical libraries from RFA will be deposited in the MLSMR and tested for bioactivities in MLSCN centers and expanded if active
  - FY05: 8 Biotechnology Resource Grants awarded
    - P41s, 3 years
  - FY06 RFA-RM-06-003

Assay Development
Technology Development
Supporting the HTS Assay Pipeline

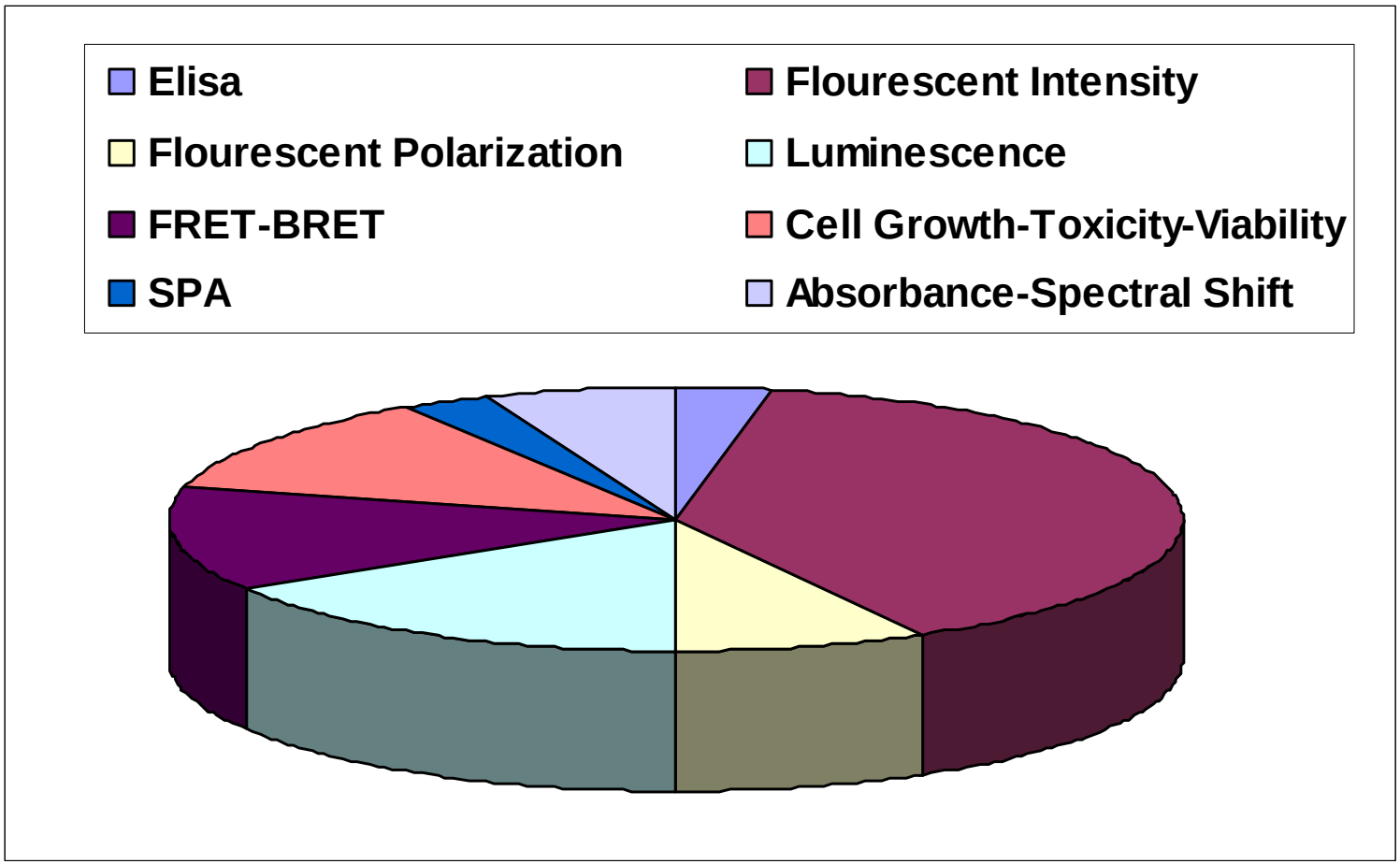
- Assay Development for High Throughput Molecular Screening (RFA-RM-06-004)
  - http://nihroadmap.nih.gov/molecularlibraries/grants.asp
  - Facilitate the development of innovative assays
  - Provide a continuously evolving stream of assays that can enter into the MLSCN via the HTS solicitation
    - Biochemical or cell-based assays of activity or interaction involving proteins and/or other biological molecules
    - Assays of cellular or molecular phenotypes
    - Modulation of expression of genes of interest, including effects on transcription, translation or RNA splicing
    - Assays involving mutant proteins associated with disease
    - Cell-based assays of cell signaling or biosynthetic pathways
  - FY04: 29 R03s awarded (1 year)
  - FY05: 38 R03s/R21s awarded (1 year)

Assay Development Awards To Date

Assay Formats

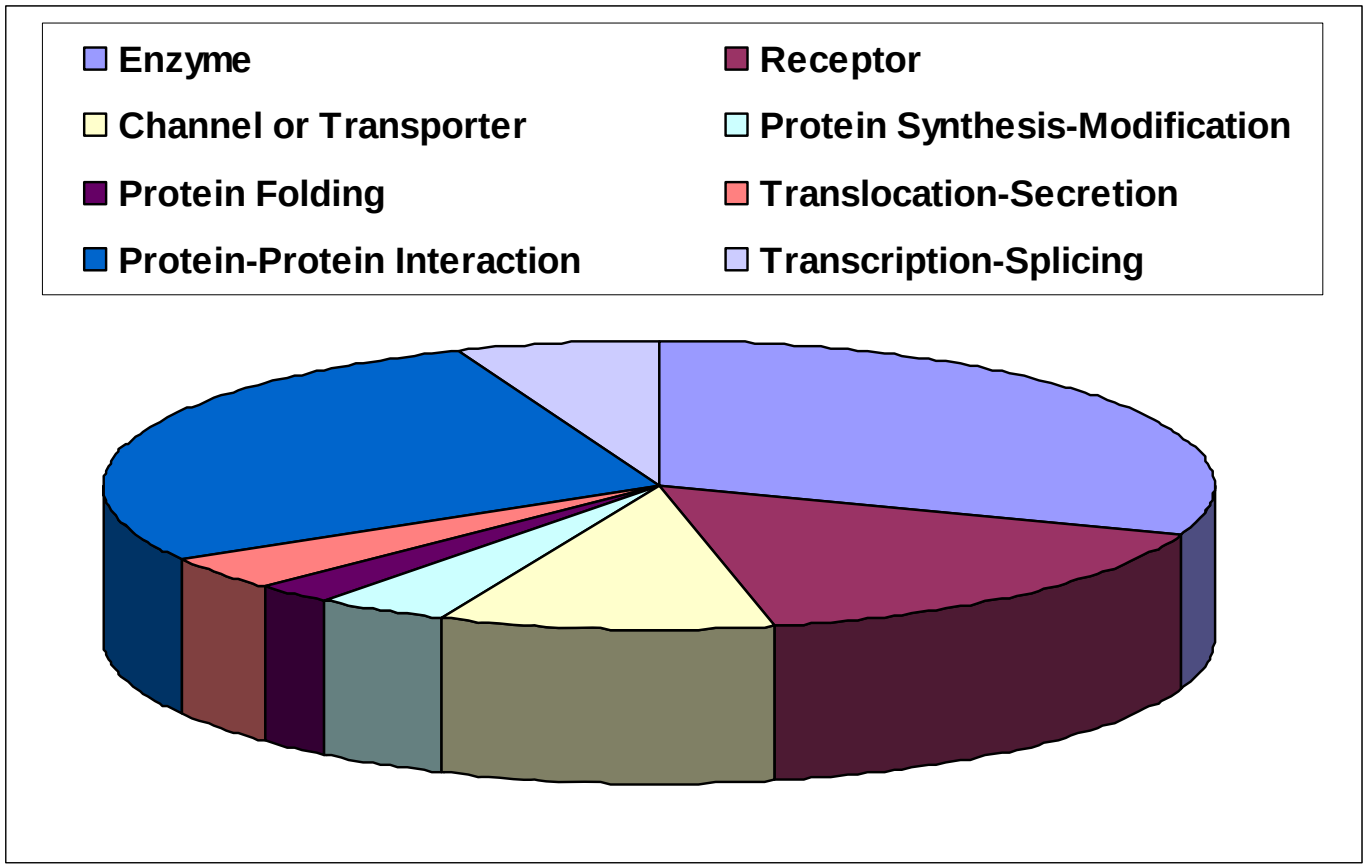


Assay Development Awards To Date
Detection Technology



Assay Development Awards To Date

Molecular Assay Targets



Chemical Diversity
Technology Development

- New Methodologies for Natural Products Chemistry (RFA-RM-05-013)
  - Rationale: Natural products have higher rate of bioactivity than novel synthesized compounds
    - most small molecule new chemical entities introduced as drugs worldwide 1981–2002 were NPs or NP-related
    - but Pharma has largely abandoned NPs in favor of synthesis because NP chemistry is slow and NP supply problematic
  - New methodologies needed to address these problems
    - Rapid isolation, purification, and identification
    - Chemical and genetic strategies for derivatization
    - Rapid isolation of genes for biosynthesis
    - Universal expression systems
  - FY05: 6 R01s awarded (3 years)

HTS Instrumentation

Technology Development

- ML Screening Instrumentation (RFA-04-020)
  - Purpose: Support for technology development in HTS instrumentation necessary to achieve MLI objectives for throughput and for diversity of biological targets and approaches
  - Research emphasis in 3 areas
    - overcoming bottlenecks in HTS production, *e.g.*, improved robotics and systems integration
    - increasing efficiency, *e.g.*, miniaturization
    - innovation in detection methodologies, *e.g.*, protein-protein association measurement, cell imaging
  - FY05: 6 Exploratory Centers awarded
    - P20s, 2 years

Predictive ADME-Toxicology
Technology Development

- Novel Preclinical Tools for Predictive ADME-Toxicology (RFA 04-023)
  - Objectives
    - Seek novel preclinical tools to improve predictive evaluation of new chemical entities
    - Improve understanding of how drug molecules are absorbed, distributed or excreted from the body
    - Understand mechanisms by which drug molecules cause toxicities
  - FY05: 5 Exploratory/Developmental Grants awarded
    - R21, 4 years

Summary of MLSCN Opportunities

The New Interface of Chemistry and Biology