MLSCN Project Team Policy on Data Sharing and IP in the MLSCN Program

Summary of the NIH MLSCN Project Team Data Sharing and IP Focus Group Meeting
October 11, 2005

The MLSCN Project Team unanimously concurred that the data sharing and Intellectual Property (IP) plan remain unchanged from that which was specified in the Molecular Libraries Screening Centers Network (MLSCN) RFA RM-04-017, released on April 21, 2004, and in the addendum, NOT-RM-04-014, released on July 22, 2004.

This proposal was developed after consultation by NIH staff with a broad range of interested parties, including scientists, university technology transfer officials, and individuals from private foundations, pharmaceutical companies and biotechnology companies. It has also been reviewed by the NIH MLSCN Project Team and the Molecular Libraries Implementation and Imaging Group (MLIIG) after considering recently made arguments by the MLSCN centers against the current data release/IP plan. The rationales for the plan are:

I Goals of Molecular Libraries Screening Centers Network:
The NIH Roadmap Molecular Libraries and Imaging Initiative is a research program designed to develop small organic molecules that can be used as chemical probes[1] to study the functions of genes, cells, and biochemical pathways, thereby providing new ways to explore the functions of major components of cells in health and disease.

a) The Molecular Libraries Screening Center Network (MLSCN) will provide biomedical researchers with scientific resources that include annotated information (activity in a wide range of biological assays) for compounds in the Small Molecule Repository (SMR), assay automation, high-throughput screening, and synthetic chemistry capabilities that can be applied to the discovery and development of innovative chemical tools.
b) All data generated by the MLSCN will be deposited promptly upon data verification[1] into PubChem. For purposes of this policy, the term "data" will include, but will not be limited to, assay descriptions, protocols and/or links to published assays implemented in the MLSCN; performance data for assays and compounds; primary data from HTS and data generated in the secondary screen (e.g., EC50s, IC50s, AC50s, counter screens); chemical structures, synthesis protocols, and/or links to published synthesis protocols for chemical analogs of hits, for probes, and the biological activity of analogues and chemical probes. PubChem is a public database; through it, the MLSCN data will be available to all researchers, in both the public and private sectors, for further use in studying biology and disease.

II Key requirements of the MLSCN data sharing and IP plan:
a) Assays and improvements in assay methods may be patented.
b) Primary screening by MLSCN:
i. Compounds from the SMR identified as "hits" from primary screening are pre-competitive.
ii. Upon verification[1] of "hit" activity for a compound, the screening results, compound structure and performance data on the "hit" must be promptly deposited in PubChem.
iii. The MLSCN centers will define data verification for an assay and receive acceptance from the NIH program managers prior to commencing work on the assay.
iv. Deposition of screening data into PubChem will not be delayed by secondary screening for "hit" investigation[1].

c) Secondary screening:

[1] See definition of terms at end of proposal

i. Results from secondary screening of initial “hits” and chemical analogs of “hits” must be promptly deposited in PubChem.

d) Synthetic chemistry and chemical probe development:
   i. Development of chemical analogs of “hits” should be made solely in order to meet the specifications of a chemical probe (see definition of chemical probe).
   ii. Commitment of synthetic chemistry resources to develop chemical analogs of “hits” is left to the discretion of the MLSCN centers and the NIH Scientific Program Managers.
   iii. 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.
   iv. Once the specifications of chemical probe are achieved, further chemical optimization ceases within the MLSCN.
   v. Following deposition in PubChem of the chemical probe data, the centers are free to pursue further chemical optimization independently of the MLSCN Program and MLSCN funds.
   vi. Screening centers may not commit synthetic chemistry resources on MLSCN assays without initial deposition of data into PubChem from the HTS primary screen and secondary screen (if applicable).

III Exceptions for two unique circumstances:
   a) First exceptional circumstance: if the 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 would be considered by the NIH MLSCN Project Team to allow early publication of the screening results. The delay in data deposition would not exceed 60 days.
   b) Second exceptional circumstance: 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 would be considered by the NIH MLSCN Project Team to allow filing of a patent. The delay in data deposition would not exceed 60 days.

IV Steps to compliance:
   a) NIH data sharing and IP plan must be made absolutely clear to centers.
   b) Justification for requirements outlined in data sharing and IP plan:
      i. Pilot phase of MLSCN designed to examine the interest of the scientific community to provide assays for screening and to utilize the resources provided in PubChem.
      ii. Scientific community will not provide assays if they see screening centers capitalizing on results based on their assays.
      iii. Screening centers should not attain an unfair advantage over competitors through government sponsored grants.
      iv. Patents produce restrictions on use of methods and materials. Limits on use of methods and materials generated by the MLSCN restrict the opportunity for these items to benefit the public good.
      v. Data sharing and IP requirements clearly defined in Molecular Libraries Screening Centers Network RFA.
   c) NIH program management will monitor Center performance, including time between receipt of assays and deposition of data in PubChem, and adherence to the data sharing and IP plan and periodically check new patent applications.
   d) Response to Center non-compliance:
      i. Redistribution of funds in next year budget.
      ii. Consider including a Determination of Special Circumstances (DEC) in competitive renewal of the MLSCN program.
      iii. Change from a cooperative agreement renewal of the MLSCN program to contracts, also consider including a DEC, when time for competitive renewal.

V Definition of terms:

a) Chemical probe: a chemical compound with activity in the primary and any secondary assays with adequate potency and aqueous solubility to be useful for in vitro (i.e., cell-based) experimentation. The specifications for a probe are likely to vary depending on the target and may need to be set jointly by the assay provider and the MLSCN Steering Committee. An example of the specifications for a chemical probe that would define an endpoint for MLSCN activities would be: expected potency of ≤ 1 uM and solubility in ≤ 1% DMSO.

b) Hit: a compound or substance whose presence in an assay modifies one or more measurable parameters producing a desirable or predictable outcome.

c) Verification: technically reproducible data points.

d) Data verification: the generation of data demonstrating a reproducible response from a specific compound or substance. The centers assigned the assay will define the method of verification which must be accepted by a NIH program manager.

e) “Hit” investigation: generation of further biological data on the compound through additional means of testing data such as secondary assays.

f) Primary screening: the initial testing of compounds to identify specific compounds whose presence in an assay modified the assay performance.

g) Secondary screening: assays designed to provide additional information on compounds of interest from their behavior in an earlier primary screen.