Harvard Institute of Chemistry and Cell Biology (ICCB)
The Initiative For Chemical Genomics (ICG)

Diversity-Oriented Synthesis (DOS):
A Platform for Discovery in Chemical Genetics

Jared Shaw – Institute Fellow (ICCB)
Co-Project Leader, Harvard Center for Methodology and Library Development (HCMLD)
jared_shaw@hms.harvard.edu

http://iccb.med.harvard.edu
http://chembank.med.harvard.edu
http://iccb.med.harvard.edu/chemistry/shaw_group/index.htm
Diverse Goals of the ICCB

- Chemistry - To bring the power of modern synthetic chemistry to bear on problems in Cell Biology.
  - Diversity-oriented, split-pool synthesis

- Biology - To increase our understanding of fundamental mechanism in cell biology.
  - Via Screening broadly for activities/phenotypes.

Information Analysis enables both of the above. (this powers our iterative approach)
Key Enabling Technologies of the ICCB

• Chemical Synthesis platform - flexible, economical, scalable, integrated,
  enables split-pool diversity-oriented synthesis (DOS)
• Screening program - economical, high throughput and flexible.
• Informatics - system that is essentially open for all.
• ChemBank - fully QC’d central repository of data, technology, how to do this, etc…

Goals
• accomplished economically, efficiently in an academic setting.
• Portable and Publicly accessible
• We are not trying to discover drugs.
A technology platform for diversity-oriented synthesis: One bead-one stock solution (Version 2.0)

1. **forward chemical genetics**
   - pin transfer to high density wells for phenotypic assays

2. **reverse chemical genetics**
   - recapture small molecules on microarrays for binding assays

**ChemBank**

(A) $\text{Si} \quad \text{i-Pr} \quad \text{i-Pr}$

(B) $\text{OMe}$

\[ \text{linker} \quad \text{HO-substrate} \]

(1) load substrate

(2) pathway development (diversity-oriented organic synthesis)

(3) library realization (encoding protocol)

(4) bead arraying

(5) compound release

(6) stock solution preparation

(7) pin transfer robot

(8) microarraying robot

(9) "hit" decoding

(10) resynthesis and maturation

(11) database entry

\[ \text{tag-1} \quad \text{tag-2} \quad \text{tag-3} \quad \text{small molecule} \]

= 500-600 $\mu$m beads 0.1 mg compound/bead

= C, Si-linker-functionalized mimotope lantern; 10 mg compound/lantern

> 5 mM stock solutions

ChemBank

384-well stock plates
A technology platform for diversity-oriented synthesis: One bead-one stock solution (Version 2.0)

1. forward chemical genetics
   pin transfer to high density wells for phenotypic assays

2. reverse chemical genetics
   recapture small molecules on microarrays for binding assays

ChemBank

(A) "hit" decoding
(10) resynthesis and maturation
(11) database entry

(1) load substrate
(2) pathway development (diversity-oriented organic synthesis)
(3) library realization (encoding protocol)
(4) bead arraying
(5) compound release
(6) stock solution preparation
(7) pin transfer robot
(8) microarraying robot

= 500-600 µm beads
0.1 mg compound/bead

= C, Si-linker-functionalized mimotope lantern; 10 mg compound/lantern

384-well stock plates > 5 mM stock solutions
How we got to version 2.0…

- 80 µm tentagel
- 100 pmol/bead
  - Not a robust/efficient Linker system
  - Arraying beads problematic
  - Enough for a **single** assay.
  - Must resynthesize every molecule!

- 600 µm polystyrene
- Minimally 50 nmol/bead
  - Silicon-based Linker system
  - HF/pyr cleavage system
  - Single Bead per well arrayer
  - Arrayed stock DMSO solutions (~ 4 mM)
  - Enough material for ‘cherry-picking’
  - Realistically, ~ 130 nmol/bead
Split-Pool Synthesis

Diversity step

Split step

Pool step

x 12

9 molecules
For Screening

Remove 3 molecules (1 each)
for screening; Carry 9 forward

6 reactions → 12 molecules
Most Libraries to Date are “Mono-skeletal”

The dihydropyran carboxamide (DHPC) library: Stavenger & Schreiber 2001

Advantages

• Short, linear synthesis-the number of reactions minimized.
• First Step is enantioselective and catalytic- changing enantiomer of catalyst is an element of diversity

Disadvantages (real and potential)

• Synthesis of building blocks detracts from the efficiency of this library
• Lack of three-dimensional diversity could bias library toward certain targets, away from others
Skeletal Diversity May Offer Greater Variety of Bioactivity from Smaller Libraries

**Skeletal Determinant:**
1) Alters Placement of Building Blocks in 3-Dimensional Space
2) Preferable if Commercially available or synthesized in <4 steps
3) 2-6 Units Necessary for $10^3 - 10^4$ compounds

**Building Block:**
1) Presents functionality to biological target
2) Must be commercially available for maximum efficiency
3) 20-50 units necessary for $10^3 - 10^4$ compounds

**Diversity Step 1:** Attach Skeletal Determinant #1 to Solid Phase

**Diversity Step 2:** Attach Building Block #1

**Diversity Step 3:** Attach Building Block #2

**Diversity Step 4:** Attach Skeletal Determinant #2

Compounds With Widely Varied Core Structures Originating from Common Precursors
Can a linear Synthetic Sequence Produce Molecules with Different 3-dimensional Structures?

Three-dimensional representations of the 4 possible skeletal arrays
Pathway Development

\[ \text{Pathway Diagram} \]

85% Yield

>95:5 diastereoselection

Trace

Trace

NOT OBSERVED

\[ \text{Chemical Structures} \]
**Pathway Development**


Pathway Development

\[
\text{O} \quad \text{O} \\
\text{I} \quad \text{I}
\]

\[
\text{O} \quad \text{N} \quad \text{Bn} \\
\text{O} \quad \text{I} \quad \text{NH}
\]

\[
\text{NHMe} \quad \text{NHMe}
\]

\[
\text{C}_6\text{H}_6/\text{RT} \quad \text{CuI (5 mol%)/K}_2\text{CO}_3 \quad \text{NO REACTION}
\]

\[
\text{MeO} \quad \text{I}
\]

\[
\text{MeO} \\
\text{Bn} \\
\text{O} \\
\text{I} \\
\text{NH}
\]

\[
\text{>95:5 Diastereoselection}
\]

\[
\text{NO REACTION}
\]

\[
\text{MeO} \quad \text{Bn} \\
\text{NHMe} \\
\text{NHMe}
\]

\[
\text{Cul (5 mol%)/K}_2\text{CO}_3 \\
\text{CuI/CsOAc} \quad \text{DMSO/ 90 °C}
\]

\[
\text{>95% conversion}
\]
Pathway Development

\[
\text{BnNH}_2 + \overset{\text{BnNH}_2}{\text{H}} \rightarrow \overset{\text{BnNH}_2}{\text{H}}
\]

\[
\text{Solvent (reflux)} \rightarrow \overset{\text{anti}}{\text{anti}} \text{ Yield 85-90%}
\]

\[
\text{Solvent} \quad \% \text{ yield} \quad \text{anti} : \text{syn}
\]

<table>
<thead>
<tr>
<th>Solvent</th>
<th>% yield</th>
<th>anti : syn</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH\text{<em>}Cl\text{</em>}2</td>
<td>80</td>
<td>83 : 17</td>
</tr>
<tr>
<td>CH\text{Cl}_3</td>
<td>90</td>
<td>83 : 17</td>
</tr>
<tr>
<td>EtOAc</td>
<td>88</td>
<td>75 : 25</td>
</tr>
<tr>
<td>THF</td>
<td>85</td>
<td>70 : 30</td>
</tr>
<tr>
<td>EtOH</td>
<td>79</td>
<td>70 : 30</td>
</tr>
<tr>
<td>DMF</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CH\text{_}3CN</td>
<td>92</td>
<td>65 : 35</td>
</tr>
<tr>
<td>Benzene</td>
<td>94</td>
<td>75 : 25</td>
</tr>
<tr>
<td>Toluene</td>
<td>93</td>
<td>90 : 10</td>
</tr>
</tbody>
</table>

\[
\text{X} = \text{I} \text{ or Br, No Reaction}
\]

\[
\text{X} = \text{I, R} = \text{H, 95% yield}
\]

\[
\text{X} = \text{I, R} = \text{CH}_2\text{OSi(i-Pr)_3, 95% yield}
\]
Pathway Development

97% yield
80:20 Diastereoselection

>95% yield

Carboxylate

Aryl-Stabilized Enolates
Pathway Development

\[
\text{Si-O}_i\text{-Pr}_i\text{-Pr} = R
\]

<table>
<thead>
<tr>
<th>Anhydride</th>
<th>SM</th>
<th>Product (anti : syn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>—</td>
<td>&gt;95 (76 : 24)</td>
</tr>
<tr>
<td>OMe</td>
<td>16</td>
<td>84 (80 : 20)</td>
</tr>
<tr>
<td>Cl</td>
<td>&lt;5</td>
<td>&gt;95 (71 : 29)</td>
</tr>
<tr>
<td>CF₃</td>
<td>&lt;5</td>
<td>&gt;95 (69 : 31)</td>
</tr>
<tr>
<td>NO₂</td>
<td>—</td>
<td>&gt;95 (&gt;95 : 5)</td>
</tr>
</tbody>
</table>
Summary

(X = H or I)
A technology platform for diversity-oriented synthesis: One bead-one stock solution (Version 2.0)

1. forward chemical genetics
   - pin transfer to high density wells for phenotypic assays

2. reverse chemical genetics
   - recapture small molecules on microarrays for binding assays

ChemBank

384-well stock plates
> 5 mM stock solutions

(1) load substrate
(2) pathway development (diversity-oriented organic synthesis)
(3) library realization (encoding protocol)
(4) bead arraying
(5) compound release
(6) stock solution preparation
(7) pin transfer robot
(8) microarraying robot
(9) "hit" decoding
(10) resynthesis and maturation
(11) database entry

(A) = 500-600 µm beads
0.1 mg compound/bead

(B) = C, Si-linker-functionalized mimotope lantern;
10 mg compound/lantern

= tag-1
= tag-2
= tag-3
= small molecule
Bead Arrayer

Illustration By: Jim Horn

Prof. Randy King, Les Walling
Self-Contained Cleavage System
Library Formatting

Compound and Beads

Beads

Tag Cleavage & Decoding

2 DMSO Stock Plates @ 5mM

Cell Based Assays

DMF Stock Plate

Small Molecule Printing

Storage & Reference Samples

Abram Calderon, Xiaohua Li
A technology platform for diversity-oriented synthesis: One bead-one stock solution (Version 2.0)

1. **forward chemical genetics**
   - pin transfer to high density wells for phenotypic assays

2. **reverse chemical genetics**
   - recapture small molecules on microarrays for binding assays

(A) 500-600 µm beads
0.1 mg compound/bead

(B) C, Si-linker-functionalized mimotope lantern; 10 mg compound/lantern

ChemBank

**ChemBank**

1. load substrate
2. pathway development (diversity-oriented organic synthesis)
3. library realization (encoding protocol)
4. bead arraying
5. compound release
6. stock solution preparation
7. pin transfer robot
8. microarraying robot
9. "hit" decoding
10. resynthesis and maturation
11. database entry

384-well stock plates
> 5 mM stock solutions
Classical Genetics/Chemical Genetics

Classical genetics

- inactivating mutation in gene C
- activating mutation in gene C

- genetic approach
- chemical genetic approach

- cytoplasm
- nucleus

- cell division, differentiation, cell death, etc.

Chemical genetics

- colchicine
- dexamethasone
Collaborative Screening at the ICCB

patterned after forward genetics  patterned after reverse genetics

**Phenotypic Screening**
- screen collections of structurally complex and diverse small molecules in search of a particular cellular/organismal phenotype.
- ICCB hosts investigators

**Proteomic Screening**
- modify cellular/organismal function of a targeted protein with a small molecule, search broadly for the resulting phenotype.
- ICCB will distribute microarrays

---

htp cytoblot assays  htp zebrafish, worm screening-by-imaging

small molecule and protein microarrays
A technology platform for diversity-oriented synthesis:
One bead-one stock solution (Version 2.0)

1. **forward chemical genetics**
   - Pin transfer to high density wells for phenotypic assays

2. **reverse chemical genetics**
   - Recapture small molecules on microarrays for binding assays

(A) = 500-600 µm beads
0.1 mg compound/bead

(B) C, Si-linker-functionalized mimotope lantern;
10 mg compound/lantern

(1) load substrate
(4) bead arraying
(7) pin transfer robot
(10) resynthesis and maturation
(11) database entry

ChemBank

(2) pathway development (diversity-oriented organic synthesis)
(3) library realization (encoding protocol)
(5) compound release
(6) stock solution preparation

(9) "hit" decoding

(8) microarraying robot

384-well stock plates
> 5 mM stock solutions
Library Formatting

Compound and Beads → residue → Master copy → Storage & Reference Samples

Tag Cleavage & Decoding

2 DMSO Stock Plates @ 5mM
Cell Based Assays

DMF Stock Plate
Small Molecule Printing

Abram Calderon, Xiaohua Li
Typical GC run for decoding solution
Manual Decoding Process

Chemistry → Beads → GC/LCMS

Visual analysis & manual data entry → Excel files

Afferent → SDF file → ChemFinder
Disadvantages of Manual Decoding

• Visual analysis, manual data entry and maintenance of the data is slow and laborious
• Excel is not suited to storing data large amounts of decoding data
• The process is so time-consuming that generally only a small subset of compounds (those that hit in a screen) are actually decoded
Can Chemical Encoding be integrated with Library Enumeration into an Automated Process?

- Chemistry
- Beads
- GC
- Afferent
- SDF file
- Automated Decoding System
A technology platform for diversity-oriented synthesis: One bead-one stock solution (Version 2.0)

1. forward chemical genetics
   pin transfer to high density wells for phenotypic assays

2. reverse chemical genetics
   recapture small molecules on microarrays for binding assays

* = 500-600 µm beads
0.1 mg compound/bead

(A) = 500-600 µm beads
0.1 mg compound/bead

(B) = C, Si-linker-functionalized mimotope lantern;
10 mg compound/lantern

(1) load substrate
(2) pathway development (diversity-oriented organic synthesis)
(3) library realization (encoding protocol)
(4) bead arraying
(5) compound release
(6) stock solution preparation

384-well stock plates
> 5 mM stock solutions

(7) pin transfer robot
(8) microarraying robot

ChemBank
A technology platform for diversity-oriented synthesis: One bead-one stock solution (Version 2.0)

1. **forward chemical genetics**
   - Pin transfer robot
   - Pin transfer to high density wells for phenotypic assays

2. **reverse chemical genetics**
   - Microarraying robot
   - Recapture small molecules on microarrays for binding assays

- **ChemBank**
- **tag-1**
- **tag-2**
- **tag-3**

**Materials and Methods**

- **Si-O** substrate
- **Si-O** encoded library
- **384-well stock plates**
- > 5 mM stock solutions

**Key Points**

- **i-Pr**
- **Si**
- **OMe**

**Reaction Details**

- **(1) load substrate**
- **(2) pathway development**
- **(3) library realization**
- **(4) bead arraying**
- **(5) compound release**
- **(6) stock solution preparation**
- **(7) pin transfer robot**
- **(8) microarraying robot**
- **(9) “hit” decoding**
- **(10) resynthesis and maturation**
- **(11) database entry**

**Note**

- 500-600 µm beads
- 0.1 mg compound/bead

**ChemBank**

**References**

- C, Si-linker-functionalized mimotope lantern;
- 10 mg compound/lamp

**Legend**

- **A**
- **B**

**Equations**

- (1) Si-O
- (2) C
- (3) Si
- (4) Pr
Acknowledgments

ICCB Professional Staff
chemistry
Xiaohua Li
Jennifer Raggio
Leticia Castro
Kerry Pierce
Max Narovlyansky
Hua Maio
Li Lai

Screening
Caroline Shamu (head)
James Follen
David Hayes
Katrina Schulberg

Informatics
Erik Brauner (head)
Jermey Muhlich

Shaw Laboratory
Yoshiyuki Fukase
Pui-Yee Ng
Craig Masse
Maria Sanchez-Rosello
Judy Mitchell

Dr. John Tallarico

Dr. Rebecca Ward
Prof. Stuart Schreiber
Prof. Tim Mitchison

ICCB Laboratories
Clemons Laboratory
Tallarico Laboratory
Schreiber Laboratory
King Laboratory
Feng Laboratory
Mitchison Laboratory

Funding
NCI
HCMLD
NIGMS
ICG

Daylight
Chemical Information Systems, Inc.

The forefront of cheminformatics.