Combinatorial libraries: strategies and methods for ‘lead’ discovery

Alan Spivey
Department of Chemistry
University of Sheffield
Key sources of information

• WWW:
  – Diversity information pages [http://www.5z.com/divinfo/]

• Books

• Reviews
Format and scope of lecture

• What is combinatorial chemistry?
• The drug discovery process
• Approaches to combinatorial library synthesis:
  – mix and split synthesis
  – parallel synthesis
  – encoded tagging
• Library types:
  – oligomeric libraries
  – template based libraries
• Combinatorial drug discovery!
What is combinatorial chemistry?

Combinatorial chemistry is a useful tool for rapidly optimizing molecular properties, particularly ones that are difficult to design \textit{a priori}…

Nature uses a combinatorial approach to generate diverse functional macromolecules such as antibodies to recognize a vast array of antigens.
The drug discovery process

• The total cost of bringing a new drug to market is typically ~£250m (i.e. EXPENSIVE!)
• Of this, £170m is spent on DISCOVERY RESEARCH.
• This reflects the large amount of TIME involved in synthesising new compounds.
• A typical chemist can synthesise ~100 compounds a year using traditional techniques.
• SOLID PHASE ORGANIC SYNTHESIS (SPOS) and COMBINATORIAL CHEMISTRY are beginning to revolutionise this situation.
Discovery chemistry: stage 1

- **High Throughput Screening (HTS):**
  - Rapid, automated screening of compounds for specific biological activity.

- **Role of combinatorial chemistry:**
  - Very large libraries.
  - Maximum diversity libraries.
  - Mix and split libraries (& parallel synthesis).
  - Mixtures of compounds (& single compounds).
Discovery chemistry: stage 2

- **Medicinal chemistry:**
  - Systematic optimisation of molecular and physicochemical properties of lead compound

- **Role of combinatorial chemistry:**
  - Small libraries.
  - ‘Targeted/focussed’ libraries.
  - Parallel synthesis libraries.
  - Single compounds.

---

**Discovery Biology**

**Therapeutic area identified**

**Biological assay established**

**Automation of assay**

**Assay adapted for High Throughput Screening (HTS)**

**Compound HTS**

**Lead structure**

**Medicinal Chemistry**

**Combinatorial Chemistry**

**Pre-clinical candidate**
Traditional vs. combinatorial

• Traditional synthesis:

\[ A + B \rightarrow AB \]

compounds prepared one at a time, characterised and screened

• Combinatorial synthesis:

\[ \begin{array}{c}
\sim\text{A}_1 \sim\text{B}_1 + \sim\text{A}_2 \sim\text{B}_2 + \sim\text{A}_3 \sim\text{B}_3 \\
\text{reaction of 3 reagents } \text{A}_x \text{ with 3 reagents } \text{B}_y \text{ provides a library of } 3^2 \text{ (i.e. 9) compounds } \text{A}_x\text{B}_y \\
\text{introduction of a third set of 3 reagents } \text{C}_z \text{ increases the library size to } 3^3 \text{ (i.e. 27) compounds } \text{A}_x\text{B}_y\text{C}_z
\end{array} \]
Approaches to ‘combinatorial’ library synthesis

• **In vivo** - biological methods:
  – Phage display, plasmids, polysomes etc.

• **In vitro** - synthetic methods:
  – Mix and split using Solid Phase Organic Synthesis (SPOS).
    • Cleavage from the solid support following ‘mix and split’ results in complex mixtures (pools) of compounds. Screening of these mixtures yields ‘hits’ whose identity must be determined by ‘deconvolution’.
    • If screening can be performed ‘on-bead’ (i.e. ‘one-bead one-compound’ libraries) then deconvolution can be avoided.
  – Parallel synthesis using Solid Phase Organic Synthesis (SPOS).
    • Spatially separate synthesis of single compounds whose identity is uniquely defined by their location.
Mix & split synthesis: libraries of mixtures of compounds in solution

- Screening complex mixtures of compounds in solution can give false ‘hits’ due to synergistic effects.
- Identification of a compound within the mixture responsible for the ‘hit’ requires iterative deconvolution.
Mix & split synthesis: ‘one-bead one-compound’ libraries

- Requires a very sensitive screening protocol which can accommodate resin bound compounds.

- Identification of a ‘hit’ compound on (or from) a single bead (~100pm) by:
  - analytical methods e.g. Edman sequencing of peptides, MALDI-TOF MS, single bead NMR...
  - reading ‘encoding tags’ on beads.
Clark Still’s encoded tagging protocol


- Each ‘monomer’ used in the library synthesis has an associated encoded tag.

- The tags are chlorinated aromatic compound which can be analysed at sub-picomolar levels by Electron Capture Gas Chromatography (ECGC).

- Allows for hit identification at one-bead fidelity for any type of library.
Mechanism of Clark Still encoded tags


$\text{N}_2\text{O}$

10 different $\alpha,\omega$-diols ($n = 3-12$)
4 different chlorophenols ($m = 2-5$)  
40 different tags

$\text{[Rh]}$  

1) ceric ammonium nitrate (CAN)
2) $\text{OTMS}$

'Bread' using 
Electron Capture Gas Chromatography

$\text{Me}^+\text{NTMS}$
Parallel synthesis: spatial separation gives single compounds

- Suitable for the synthesis of relatively small libraries as each compound requires its own reaction ‘well’.
- Each reaction ‘well’ may be anything from a small flask to a radio-frequency tagged ‘tea bag’ to an etched region on a silicon chip!
- Once screening has identified a hit no further work is required to deduce the identity of the active compound although it is routine practice to independently verify the structure.

<table>
<thead>
<tr>
<th>Split ± cleavage from resin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screen each ‘position’</td>
</tr>
<tr>
<td>'Hit'</td>
</tr>
</tbody>
</table>

Spatial location of hit defines its identity

Identity of hit

Spatially separated libraries
Keeping track of ‘tea-bag’ parallel synthesis: Irori radio-frequency tagging

- http://www.irori.com/
There are essentially two strategically distinct types of library at the molecular level:

- **Oligomeric**
  - peptides/peptoids
  - oligonucleotides
  - oligosaccharides
  - unnatural oligomers
  - polyaromatics

- **Template based**
  - drug-like molecules
  - natural product-like molecules
  - heterocycle based molecules
Balasubramanian’s peptide library

- Protein tyrosine phosphatase substrate library (oligomeric).
- **Library synthesis:**

  1. Fmoc-Aaa-OH, PyBOP, HOBt, DIPEA
  2. piperidine

  **mix and split with encoding**

  encoding via dummy peptide sequence incorporating glycine in place of phosphotyrosine

  one-bead one-peptide encoded library
Balasubramanian’s peptide library

- Protein tyrosine phosphatase substrate library (oligomeric).
- Library screening:

  1) leukocyte antigen receptor protein tyrosine phosphatase (*PTP*)
  2) α-chymotrypsin
  3) fluorescent labelling of *N*-terminus with carboxyfluorescein:

Identity of PTP substrate sequences

"preference for at least two acidic residues in variable positions and a glutamic acid residue at *X*₁aa"
Beck-Sickinger’s cyclic peptide library

- Neuropeptide Y analogue library (oligomeric).
- **Library synthesis:**

```
1) DFOH
PPh₃, DEAD Mitsunobu
parallel synthesis
2) Fmoc-Aaa-OH, DIC, HOBt
3) piperidine

Xaa-Y-P-S-K-Xaa-R-Q-R

spatially separated solution library

1) LiOH
2) TBTU, HOBt, DIPEA
```
Beck-Sickinger’s cyclic peptide library

- Neuropeptide Y analogue library (oligomeric).
- **Library screening:**

```
Xaa-Y-P-S-K-Xaa-R-Q-R
```

spatially separated solution library

competitive binding assay in solution with radiolabelled neuropeptide Y

```
Hits
```

identity defined by spatial location

"weak competitive binding at μM level by range of derivatives"
Oligonucleotide libraries

- These oligomeric libraries are generally prepared using *in-vivo* ‘biological’ methods and screened using **Systematic Evolution of Ligands by Exponential enrichment** (SELEX) procedures.

- e.g. The discovery of very high-affinity RNA and DNA ligands to human IgE which inhibit binding to the Fcε receptor I.

- Wiegand *J. Immunology* 1996, 157, 221 (and references therein).
Unnatural backbone oligomer libraries

**Backbone**

- Oligocarbamate
- Oligo-phosphodiester
- Vinylogous sulfonamidopeptide
- Vinylogous sulfonamidopeptide

**Monomers**

- Fmoc \( R \) \( \text{NH} \) \( \text{O} \) \( \text{O} \) \( \text{P} \) \( \text{NP} \)
- \( \text{NC} \) \( \text{O} \) \( \text{P} \) \( \text{O} \) \( \text{DMTr} \)
- Boc \( R \) \( \text{NH} \) \( \text{SO}_2 Cl \)
- \( \text{Cl} \) + \( R-\text{NH}_2 \)
Schultz’s purine library

- Kinase inhibitor library (template based).
- **Library synthesis:**

PS resin with acid labile alcohol linker pre-functionalised with hydroxyethyl group

parallel synthesis

spatially separated solution library

Nucleophilic substitution (C-2)

TFA
Schultz’s purine library

- Kinase inhibitor library (template based).
- **Library screening:**

Screen each position for inhibition of h-CDK2-cyclin A kinase complex. Hits defined by spatial location.

- 3 & 4 substituted anilines & benzylamines
- Ala, Val or Ile derived amino alcohols
- Small alkyl e.g. Me

Spatially separated solution library

- $R_1$-$NH$
- $R_2$-$OH$
- $R_3$
- $R_4$
Schreiber’s ‘natural product’ library

- Protein epitope binding library (template based).
- Library synthesis:

  ![Chemical structures](image)

  - *PS* with photolabile amide linker
  - *mix and split with Clark Still encoding*
  - *one-bead one-compound encoded library*
Schreiber’s ‘natural product’ library

- Protein epitope binding library (template based).
- Library screening:

One-bead one-compound encoded library

"...several members of this library activate a reporter gene in mink lung cells!"
Nicolaou’s sarcodictyin library

- Tubulin-microtubule disruptant library (template based).
- Library synthesis:

![Diagram of library synthesis]

\[ \begin{align*}
\alpha,\omega\text{-dil functionalised hydroxymethyl-PS} & \quad \xrightarrow{\text{PPTS, CH}_2\text{Cl}_2} \\
\text{spatially separated library in IRORI 'tea bags'} & \quad \xrightarrow{\text{parallel synthesis}} \\
\end{align*} \]
Nicolaou’s sarcodictyin library

- Tubulin-microtubule disruptant library (template based).
- **Library screening:**

  ![Chemical structure](image)

  - Spatially separated library in IRORI 'tea bags'
  - Screen for induction of tubulin polymerisation and cytotoxicity with ovarian cancer cells
  - Hits identity defined by spatial location

  - Side chain crucial
  - Both N's important
  - Ketals tolerated
  - Esters preferred over amides; reduction to alcohol not tolerated
DeWitt’s quinolone library

- Ciprofloxazin analogue library (template based).
- **Library synthesis:**

![Chemical diagram]

**Parallel synthesis**

1. **Wang resin**
   - DMAP
toluene
   - Et-ester
2. **1)** MeO
   - OMe
   - THF
3. **2)** $R_1$-NH$_2$
4. **3)** $\Delta$
5. **$R_2$-NH$_2$**
   - NMP, $\Delta$
6. **TFA**

**Spatially separated library in DIVERSOMER reaction tubes**
DeWitt’s quinolone library

- Ciprofloxazin analogue library (template based).
- **Library screening:**

  ![Spatially separated library in DIVERSOMER reaction tubes](image)

Screen for gyrase inhibition in solution

Hits

Identity defined by spatial location

![Ciprofloxazin](image)
Gallop’s mercaptoacetyl proline library

- **Angiotensin Converting Enzyme (ACE) inhibitor library (template based).**
- **Gallop J. Am. Chem. Soc. 1995, 117, 7029.**
- **Library synthesis:**

  - 1) 20% piperidine, DMF
  - 2) ArCHO
  - 3) CH(OMe)₃
  - 4) Ac₂O, DIPEA

  - **Mix and split**
    - AgNO₃, Et₃N
    - MeCN

  - 1) AcS~R₂~COCl
  - 2) TFA
  - 3) ethylenediamine

  - **480 member pool of compounds**
Gallop’s mercaptoacetyl proline library

- Angiotensin Converting Enzyme (ACE) inhibitor library (template based).
- **Library screening:**

\[
\begin{array}{c}
\text{HO} \\
\text{O} \\
\text{N} \\
\text{Ar} \\
\text{Z} \\
\text{R}_1 \\
\text{O} \\
\text{R}_2 \\
\text{SH}
\end{array}
\]

480 member pool of compounds

**Screened for ACE inhibition**

**DECONVOLUTION**

VIA FOUR ITERATIONS OF SUB-LIBRARY RE-SYNTHESIS AND SCREENING

\[
\begin{array}{c}
\text{HS} \\
\text{Me} \\
\text{HO}_2\text{C} \\
\text{N} \\
\text{CO}_2\text{Me}
\end{array}
\]

\[K_i \sim 160\text{pM}\]

3 x more potent than captopril
Ellman’s benzodiazepine library

- Benzodiazepine library (template based).
- Library synthesis:

\[
\begin{align*}
\text{NH}_2 \quad &\text{aminomethyl-PS} \\
\rightarrow &\quad \text{NH}_2 \\
\end{align*}
\]

**parallel synthesis**

\[
\text{NH}_2 \quad \text{SnMe}_3 \text{Si} \\
\rightarrow &\quad \text{NH}_2 \quad \text{SnMe}_3 \text{Si} \\
\end{align*}
\]

**spatially separated solution library**

1) \( \text{R}_1 \text{COCl, Pd(0)} \)
2) TFA

1) \( \text{R}_1 \text{Li} \)
2) \( \text{R}_1 \text{Br, DMF} \)
3) 5% AcOH, \( \Delta \)
Ellman’s benzodiazepine library

- Benzodiazepine library (template based).
- **Library screening:**

  ![Chemical structure](image)

  screening of this library was not reported because it pertained that some of the compounds in the library still contained silicon due to an anomalous cleavage mechanism which was particularly troublesome when $R_1$ was an electron withdrawing substituent.
Summary

• What is combinatorial chemistry?
• The drug discovery process
• Approaches to combinatorial library synthesis:
  – mix and split synthesis
  – parallel synthesis
  – encoded tagging
• Library types:
  – oligomeric libraries
  – template based libraries
• Combinatorial drug discovery.