Introduction to Combinatorial Chemistry

Sanghee Kim
School of Pharmacy,
Seoul National University

Contents

1. History of Combinatorial Chemistry
2. Introduction of Chemical Library
3. Principle of Combinatorial Chemistry
1. History of Combinatorial Chemistry

Although combinatorial chemistry has only really been taken up by industry since the 1990s, its roots can be seen as far back as the 1960s when a researcher at Rockefeller University, Bruce Merrifield, started investigating the solid-state synthesis of peptides.

- Bruce Merrifield won the Nobel prize in chemistry in 1984 for his work on solid-phase synthesis.
- During this time, automated peptide synthesizer technology was in its infancy, and the preparation of individual peptides was a challenge.
Combinatorial chemistry was first conceived about 24 years ago - although it wasn't called that until the early 1990s. Initially, the field focused primarily on the synthesis of peptide and oligonucleotide libraries.

- **H. Mario Geysen**, distinguished research scientist at Glaxo Wellcome Inc., helped jump-start the field in 1984 when his group developed a technique for synthesizing peptides on pin-shaped solid supports. (currently, a faculty of Univ. of Virginia, Chemistry Department)
- At the Coronado conference, Geysen reported on his group’s recent development of an encoding strategy in which molecular tags are attached to beads or linker groups used in solid-phase synthesis. After the products have been assayed, the tags are cleaved and determined by mass spectrometry (MS) to identify potential lead compounds.

Another early pioneer was **Árpád Furka** who introduced the commonly used split-and-pool methods.

- **Dr. Árpád Furka** is considered to be one of the fathers of combinatorial synthesis publishing in 1982 the first split-mix synthesis work in the area of peptide synthesis. Dr. Furka has a website dedicated to combinatorial chemistry covering several historic aspects of combinatorial chemistry, including online links to his original split-mix synthesis article published in 1982.
- [http://szerves.chem.elte.hu/furka/](http://szerves.chem.elte.hu/furka/)
In 1985 Richard Houghten introduced the “tea-bag” method for rapid multiple peptide synthesis. These and other advances in manual multiple-peptide synthesis fed the beginnings of a wave of rapid bioassays based on the developing area of molecular biology.

“Tea-bag method”
Polyethylene bag with fine holes, similar to real tea-bag, are filled with resins and each bag is put in the different reaction vessels to carry out amino acid coupling reaction. After reactions, all the bags are collected and processed together for protecting group removing and washing resins to reduce the amount of time and efforts. In this method, the bag takes the role of filter and preventing resin mixing between reactions, and by labeling each bag, the synthesized peptide structure can be identified. About 100 different peptides in 500 micromol quantity could be synthesized by this method, which demonstrate a practical approach to parallel synthesis despite the fact that synthesized peptide number is moderate.

Comparatively few organic chemists undertook the preparation of ordinary organic substances on solid phases because the work is rather more complex when applied to non-oligomeric substances caused by greater variety of reactants and conditions required, and this work at first failed to develop a significant following. Solid phase organic chemistry was also comparatively underdeveloped and this held back the field. This changed in dramatic fashion after the publication of Bunin and Ellman’s seminal work on solid phase organic synthesis (SPOS) of arrays of 1,4-benzodiazepine-2-ones in 1992. Soon other laboratories published related work on this ring system, and work on other drug-like molecules followed in rapid order and the race was on.
From a historical perspective, the research efforts made in classical combinatorial chemistry can be briefly outlined in three phases:

- **The 1st Phase**: In the early 1990s, the initial efforts in the combinatorial chemistry arena were driven by the improvements made in high-throughput screening (HTS) technologies. This led to a demand for access to a large set of compounds for biological screening. To keep up with this growing demand, chemists were under constant pressure to produce compounds in vast numbers for screening purposes. For practical reasons, the molecules in the first phase were simple peptides (or peptide-like) and lacked the structural complexity commonly found in modern organic synthesis literature.

- **The 2nd Phase**: In the late 1990s, when chemists became aware that it is not just about numbers; but something was missing in compounds produced in a combinatorial fashion. Emphasis was thus shifted towards **quality** rather than quantity.

- **The 3rd Phase**: As the scientific community moved into the post-genomic chemical biology age, there was a growing demand in understanding the role of newly discovered proteins and their interactions with other bio-macromolecules (i.e. other proteins and DNA or RNA). For example, the early goals of the biomedical research community were centered on the identification of small-molecule ligands for biological targets, such as G-protein-coupled receptors (GPCRs) and enzymes. However, the current challenges are moving in the direction of understanding bio-macromolecular (i.e. protein-protein, protein-DNA/RNA) interactions and how small molecules could be utilized as useful chemical probes in systematic dissection of these interactions. By no means will this be a trivial undertaking! The development of biological assays towards understanding biomacromolecular interactions is equally challenging as the need for having access to useful small molecule chemical probes.
### Chemical Library

A **chemical library** or compound library is a collection of stored chemicals usually used ultimately in **high-throughput screening**.

- The chemical library can consist in simple terms of a series of stored chemicals.

- Each chemical has associated information stored in some kind of database with information such as the **chemical structure, purity, quantity, and physiochemical characteristics** of the compound.
Chemical Library in Drug Discovery

Development of the Hit Rate

Hit Rate: Proportion of test compounds versus commercialized products

Chemical Library

Current Opinion in Chemical Biology 2003, 7, 331
**Sources for Chemical Library**

- Parallel & Combinatorial Chemistry
- Random Synthetic Compound Collections
- Tailor-made Synthetic Compound Collections
- Peptoids
- Peptides
- Peptidomimetics
- Natural Products

**The Real Process of Drug Development**

[Diagram flowchart showing the real process of drug development including chemical library, lead identification, lead optimization, clinical development, and marketing.]
Types of Combinatorial Library in Drug Discovery

- Identification of a biological target
- Development of an assay
- High-throughput screening
  - Random Library
  - Hit
  - Combinatorial Chemistry
  - Lead-Structure
  - Focused Library
  - Lead-Optimisation
  - Development

Types of Combinatorial Library in Drug Discovery

**Random Libraries**
- multiple libraries
- many targets
- highly diverse
- mixtures
- > 5,000 compounds
- solid phase synthesis
- non-purified compounds
- on bead screening, if possible

**Focused or targeted Libraries**
- Template- scaffold library
- one target
- high structural similarity
- single compounds
- < 5,000 compounds
- synthesis in solution, solid phase
- pure compounds
- screening in solution
Chemical Library in Drug Discovery

Combinatorial Synthesis

Figure 12. (A) In general, in a combinatorial synthesis one starting material A reacts with one reactant B resulting in one product AB. (B) In a combinatorial reaction different building blocks of type A (A₁ – Aₙ) are reacted simultaneously with different building blocks of type B (B₁ – Bₙ) according to combinatorial principles, i.e. each starting material A reacts separately with all reactants B resulting in a combinatorial library A₁-B₁ – Aₙ-Bₙ.
Principles of Combinatorial Chemistry

- The basic principle of combinatorial chemistry is to prepare a large number of different compounds at the same time. Instead of synthesizing compounds in a conventional one-at-a-time manner.

- The characteristic of combinatorial synthesis is that different compounds are generated simultaneously under identical reaction conditions in a systematic manner, so that ideally the products of all possible combinations of a given set of starting materials (termed building blocks) will be obtained at once.

- The collection of these finally synthesized compounds is referred to as a combinatorial library.

Combinatorial Chemistry: The Game with the Large Numbers

![Diagram of a molecule structure]

N: numbers of different substituents on each position
Z: number of possible compounds

\[ Z = \frac{1}{2} N^6 \]

<table>
<thead>
<tr>
<th>N</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>50,000</td>
</tr>
<tr>
<td>20</td>
<td>1,600,000</td>
</tr>
<tr>
<td>30</td>
<td>12,150,000</td>
</tr>
</tbody>
</table>
Combinatorial Chemistry: The Game with the Large Numbers

**Variation of a Lead Structure: Small Changes with Big Effects**

![Chemical structures of Norfloxacin and Ciprofloxacin](image)

**Challenge:** how to find the best variation?

- 15 substituents per 7 positions
- 170,859,375 compounds

Combinatorial Chemistry: The Game with the Large Numbers

**Challenge:** how to find the best variation?

- 15 substituents per 7 positions
- 170,859,375 compounds

*Structure-Activity Relationships*

- NH₂, CF₃, F, Cl
- OH, NH₂, OCH₃, OCN
- No variation allowed
- 3-OH or prodrug, eg.,乙酸

*Possible Variations*
Synthetic Methodology for Combinatorial Library Construction

➢ Solid-Phase Organic Synthesis
   The compound library have been synthesized
   *on solid phase such as resin bead, pins, or chips*

➢ Solution-Phase Organic Synthesis
   The compound library have been synthesized
   *in solvent in the reaction flask*