DNA-Templated Organic Synthesis

Matthew Coulter
Organic-Biological Seminar
Department of Chemistry
University of Toronto
February 4, 2008
DNA-Templated Organic Synthesis

David R. Liu
Professor of Chemistry and Chemical Biology
Harvard University

• Small molecule library synthesis

• DNA assisted reaction discovery
DNA-Templated Synthesis

DNA → Transcription → Translation → Protein

DNA → Translation → Synthetic compound
## DNA-Templated Organic Synthesis

<table>
<thead>
<tr>
<th>Conventional Screening Approach</th>
<th>DNA-Templated-Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>• spatial separation generally required</td>
<td>• All molecules synthesized in one pot</td>
</tr>
<tr>
<td>• each compound analyzed individually</td>
<td>• All molecules analyzed simultaneously</td>
</tr>
<tr>
<td>• adequate material required for screen</td>
<td>• fmol - nmol scale</td>
</tr>
<tr>
<td>• workload increases with sample size</td>
<td>• workload does not scale with sample size</td>
</tr>
</tbody>
</table>
|   • Synthetic compound library size 10^9 – 10^{13}  
   • Modern high-throughput screening facility  
   • 9 – 9000 years | • Tools: molecular biology |
| • Tools: MS, NMR, HPLC etc. | |

Can DNA Encode Chemical Reactions?

Strategy: Link nucleophile and electrophile to separate strands of complementary DNA

A Sequence Specific Synthetic Reaction

- Bond formation occurs sequence specifically
- Very low background (non-templated) reactivity due to very low reagent concentrations
- DNA encoded reaction

Multiple Reaction Types Supported

All substrates tested reacted sequence-specifically:

Scope of DNA-Templated Reactivity

Many standard reactions are encoded sequence specifically at nM concentrations of reactants:

**Reductive amination:**

\[
\text{NH}_2^- + \text{PhCHO} \xrightarrow{\text{NaBH}_3\text{CN}, \text{pH 6.0, 1.5 h}} \text{PhNH}_2^- \]

81%

**Amide bond formation:**

\[
\text{O-CH}_2\text{NH}_2^- + \text{PhNHCOOH} \xrightarrow{\text{EDC-NHS}, \text{pH 6.0, 16 h}} \text{PhNHCOO}^- \]

59%

**Henry reaction:**

\[
\text{NHCHO} + \text{PhNHCOCH}_2\text{NO}_2 \xrightarrow{\text{pH 8.5, 12 h}} \text{PhNHCOCH}_2\text{NO}_2\text{O}^- \]

45%

Scope of DNA-Templated Reactivity

**Nitro-Michael Reaction:**

\[
\begin{align*}
\text{Nitro-} & \quad + \\
\text{Michael:} & \quad \text{pH 8.0, 1.5 h}
\end{align*}
\]

**Wittig Reaction:**

\[
\begin{align*}
\text{Wittig} & \quad + \\
\text{Reaction:} & \quad \text{pH 8.0, 1.5 h}
\end{align*}
\]

**1,3 dipolar cycloaddition:**

\[
\begin{align*}
\text{1,3 dipolar} & \quad + \\
\text{cycloaddition:} & \quad \text{pH 7.5, 22 h}
\end{align*}
\]

**Heck Reaction:**

\[
\begin{align*}
\text{Heck} & \quad + \\
\text{Reaction:} & \quad \text{pH 5.0, 2 h}
\end{align*}
\]

A Synthetic Code

- A number of reactions are encoded sequence specifically
- Reactivity is controlled by hybridization of complimentary DNA
- Sequence specificity: supports the faithful translation of a particular nucleic acid sequence to a given amino acid sequence
- Ribosomes decipher the genetic code during protein synthesis
- Input – output in templated synthesis
Applications to Small Molecule Library Synthesis

- Synthesize and evaluate a programmed small molecule library
  - *Multi-step DNA-templated synthesis*
  - *Select a single molecule based on a particular function*
  - *Identify member based on its associated DNA sequence*
Affinity Purification

- Biotin binds the proteins avidin and streptavidin very tightly ($K_d \sim 10^{-15}$)
- Avidin-biotin interaction used in affinity purification
Selection and Identification of a Single Target

Proof of Principle: Translation, Selection and Amplification of a Synthetic Library

1,025 starting materials

1,025 reagents

1,025 templated products
1,050,625 non-templated products

Selection and Identification of a Single Target

1,025 templated products of 1,050,625 non-templated products

1) In vitro selection - avidin beads
2) PCR amplification

Polymerase Chain Reaction

- Kary Mullis
- 1993 Nobel Prize in Chemistry
- A method of amplifying a sequence of DNA
- 20 cycles: $1 \times 10^6$ amplification of target sequence

Exponential amplification of target sequence

Biological Specificity: Selective Enrichment

1,025 templated products of 1,050,625 non-templated products

1) In vitro selection (avidin)
2) PCR amplification

Amplified DNA of selected molecules

Characterize:
- Restriction digestion
- Sequencing

Biotin linked product enriched 1000 fold after just one round of selection

Multistep DNA-templated synthesis

Multistep DNA-templated synthesis

Product self-elutes

1 – 5%

A DNA-Templated Library of Synthetic Molecules

Programmed Synthesis: A DNA Templated Library

65 DNA templates coding for 65 templated products were synthesized:

64 templated macrocycles

1 positive control macrocycle

A DNA-Templated Library

A DNA-Templated Library

- 65 macrocycles synthesized simultaneously in one pot
  - Presence of templated intermediates confirmed after each step

- Each completed macrocycle associated with a molecule of DNA which
  - Directed sequence specific synthesis
  - Contains amplifiable sequence which identifies molecule

Functional Selection and Identification of a One Macrocycle

- Entire library subjected to two rounds of selection for binding to carbonic anhydrase
- Selected DNA amplified via PCR
- Digested with restriction endonuclease which cleaves only positive control template

- One molecule selected based on function
- No spatial separation or segregation of library members required
- Only small amounts of material required: PCR amplification

DNA Assisted Reaction Discovery
Reaction discovery in organic synthesis generally begins with a targeted transformation
  • Optimization follows initial discovery or hit

An un-biased, one-pot, high throughput probe of reactivity between many functional groups under a given set of reaction conditions is a challenge in organic synthesis

DNA-Assisted Reaction Discovery

**Criteria for a reaction discovery system**

- Organize many reagents into defined substrate pairs in one pot
- Facile separation of reactive from unreactive substrate pairs
- Identification of reactive substrate combinations

**DNA-Templated Synthesis**

- Very low reagent concentrations and sequence specific annealing provide control over reactivity
- Bulk selection and purification strategies
- DNA Microarrays

DNA Microarrays

1) Hybridize mixture of labeled unknown DNA to chip

2) Wash away non-binding DNA

3) Identify unknown DNA based on hybridization to chip

• Often used in genomics studies
• Each spot contains DNA of unique, known sequence

Substrates for Reaction Discovery

Pool A x Pool B
- one pot
- one condition
- all combinations
- identify bond forming pairs

Adapting Templated Synthesis to DNA Microarrays

Library of templates encoding all possible pool A x pool B combinations:

Individual templates code for one substrate combination:

Adapting Templated Synthesis to DNA Microarrays

Pool A $\times$ Pool B
All combinations

Conditions

Disulfide reduction

Bond formation

Pool A $\times$ Pool B
All combinations

Disulfide reduction

Subject amplified DNA to microarray analysis

Reaction Discovery: A New Oxidative Coupling

Pool A $\times$ Pool B

2 pmol total material

$\text{Na}_2\text{PdCl}_4$ in $\text{H}_2\text{O}$, pH 7.0

A5 + B5

A Novel Pd Catalyzed Oxidative Coupling

Multi-milligram scale-up:

\[
\begin{align*}
\text{H} & \quad \text{Na}_2\text{PdCl}_4 \ (1.0 \text{ eq.}) \\
\text{H}_2\text{O}, \ 25^\circ\text{C}, \ 15\text{h} & \quad \rightarrow \\
& \quad \text{isolated yield } 86\%
\end{align*}
\]

This reaction was further developed in the synthesis of \(\alpha,\beta\)-Unsaturated ketones:

\[
\begin{align*}
\text{R}_2\text{N} & \quad \text{Na}_2\text{PdCl}_4 \cdot 3\text{H}_2\text{O} \ (15 \text{ mol\%}) \\
\text{R}_4 & \quad \text{CuCl}_2 \cdot \text{H}_2\text{O} \ (20 \text{ mol\%}) \\
\text{H}_2\text{O}/\text{MeCN} \ (5:1) \ (\text{O}_2 \ 1 \text{ atm}) & \quad \rightarrow \\
\text{R}_2\text{N} & \quad \text{isolated yield } 53 - 80\%
\end{align*}
\]

Proposed mechanism:
Summary

- DNA-templated organic synthesis enables sequence specific multi-step synthesis
  - One pot synthesis and selection of a library of programmed macrocycles

- Selected molecules can be identified based on associated DNA

- Many random combinations of substrates can be assessed for reactivity simultaneously

- Development of new synthetic methodology
Coded Synthesis

*Nature’s templated synthesis:*

- Application of coded reactivity to chemical synthesis
- New tools for exploring organic chemistry
Acknowledgements

- Professor Vy M. Dong
- The Dong Group Members