Characterization of guide RNAs for CRISPR Applications

Jean-Noël Lemercier
Disclosure

• I am an employee and shareholder of Editas Medicine
What is CRISPR?

CRISPR is a gene editing technology derived from the bacterial adaptive immune system that can revise, remove, and replace genes in a highly targeted manner.

In a therapeutic context, CRISPR uses a combination of 2 types of molecules to edit disease-related genes:

- **RNA**: can be designed to guide a nuclease to any DNA sequence.

- **Cas Nuclease**: the gene editor.

- **Guide RNA**: helps the nuclease find the right place to edit.

Specificity: CRISPR’s ability to only edit intended DNA targets and avoid off-target editing. Achieving high levels of specificity requires the right combination of nuclease and guide RNA.
A CRISPR Application for ex-vivo Cell-Based Medicine

+ 

Ribonucleoprotein (RNP), or RNA-guided endonuclease

Electroporation

© 2023 Editas Medicine
Editas Proprietary Engineered AsCas12a CRISPR Nuclease

• **AsCas12a** is highly specific primarily due to its DNA target engagement mechanism that is distinct from SpCas9

*Note: Matched target sites are 20mer protospacer sequences within the human genome with compatible PAMs for different nucleases (20-Ns)

1 Editing specificity from Gotta et al. Cold Spring Harbor 2019; 2 Editing efficiency from De Dreuzy et al, ASH 2019
Cas12a and Cas9 Require Different gRNAs

Detailed Characterization of the 2 types of gRNAs to follow

• Shorter 41mer gRNA
• Targeting sequence on the 3’-terminus of molecule

• Longer 100mer gRNA
• Targeting sequence on the 5’-terminus of molecule

Oligonucleotide Synthesis Process

- Stepwise addition of nucleotides at the 5' terminus
- Each addition is a cycle made of 3 main steps + final cleavage and deprotection
Effect of Coupling Efficiency on Yield

\[ \text{Yield} \sim (\% \text{ coupling efficiency})^n \]
where \( n = \# \text{ of nucleotide additions} \)

\[(0.99)^{40} = 67\% \text{ theoretical crude yield} \]
\[(0.99)^{99} = 37\% \text{ theoretical crude yield} \]

Longer RNA = Lower Yield
## Oligonucleotide Synthesis Impurities

<table>
<thead>
<tr>
<th>Impurity</th>
<th>Side Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truncations</td>
<td>n-x (n-1, n-2, etc.) depurination</td>
</tr>
<tr>
<td>Protecting groups</td>
<td>Failure to remove protecting groups</td>
</tr>
<tr>
<td>Synthesis errors</td>
<td>Double coupling</td>
</tr>
<tr>
<td></td>
<td>Missed base</td>
</tr>
<tr>
<td></td>
<td>Premature capping</td>
</tr>
<tr>
<td>Amidite impurities</td>
<td>Side reactions</td>
</tr>
</tbody>
</table>

1. Each synthesis step has the potential to introduce impurities
2. The number and complexity of impurities increase with gRNA length
3. Solid-phase RNA synthesis is carried out in a 3’ to 5’ direction

Sequence fidelity decays toward 5’ terminus
Purity Characterization – Liquid Chromatography

- Reverse-phase ion-pairing chromatography (RP-IP-UPLC) (industry standard)
- HFIP in mobile phase (chromatographic resolution and mass spec compatibility)
- UV Detection at 260 nm
Analytics: Guide RNA Crudes vs. Length

Waters® Acquity Oligonucleotide BEH C18, 130Å, 1.7μm, 2.1 mm x 100 mm
Mobile phase A: 95 mM HFIP, 14 mM TEA, 1%ACN
Mobile phase B: ACN
Flow rate: 0.2 ml/min
Column temperature: 50 °C
The Reality of the Data: Purity Determination by LC-UV is Imperfect

Full-Length Product and Co-eluting impurities

LC-UV Quantitation
Deletion sequence impurities
Anticipated level 0.1-3.0%

LC-UV Quantitation
Addition sequence impurities
Anticipated level 0.1-2.0%

LC-UV-MS is required to obtain a full overview of all gRNA related substances
LC-UV-MS purity < LC-UV Purity

Identity: Determination by Mass Spectrometry

Biomolecule Characterization via ESI-MS detection

Electrospray ionization
  - Revolutionary for biomolecule MS
  - Nobel prize in Chemistry 2002
    o John Fenn and Koichi Tanaka

\[ M_{\text{meas}} = 18079.8 \text{ Da} \]
\[ M_{\text{calc}} = 18079.3 \text{ Da} \]

Deconvolution
Preparative-Scale Purification of Crude by Reverse Phase Chromatography

Early-eluting Fraction

Purest Fraction

Late-eluting Fraction

Absorbance (mAU)

Time (min)

260 nm Trace
Cas12a and Cas9 Guide RNA Fractions

Cas12a gRNA

1 Early
2 Purest
3 Late

Absorbance
Time (min)

Cas9 gRNA

1 Early
2 Purest
3 Late

Absorbance
Time (min)

LC-UV Chromatograms at 260 nm
Guide RNA Purest Fractions: TIC vs EIC (Full-Length Product)

- Total Ion Chromatogram
- Extracted ion chromatogram for full-length product (FLP)
Guide RNA Purest Fraction: Cas9 Impurity Mining

- n-9
- n-5
- n-6
- FLP
- m+41
- m+41+Na
- m+41-rU
- n+1
- n+2
- n+3

Mass (Da) vs. Time (min)

%
Cas9 Guide RNA Purest Fraction LC/MS Analysis

Average mass spectrum

Deconvolution

TIC

Time (min)

Mass (Da)
Cas9 Guide RNA Purest Fraction Deconvoluted Mass Spectrum

![Graph showing mass spectrum with peaks labeled n-9, n-6, n-5, n-4, n-3, n-2, n-1's, n+rG, n+rU, and n+2. Peaks are indicated at specific mass values including 28000, 29000, 30000, 31000, and 32000 Da. Arrows indicate in-source depurination and metal adducts.]

© 2023 Editas Medicine
<table>
<thead>
<tr>
<th>Mass Delta from FLP (Da)</th>
<th>Identity</th>
<th>Peak Height</th>
<th>%Peak Height</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>-4288.8</td>
<td>?</td>
<td>180000</td>
<td>0.7</td>
<td>?</td>
</tr>
<tr>
<td>-4021.6</td>
<td>?</td>
<td>202000</td>
<td>0.8</td>
<td>?</td>
</tr>
<tr>
<td>-3574.5</td>
<td>n-12</td>
<td>224000</td>
<td>0.9</td>
<td>5’-rArU.............................................-3’</td>
</tr>
<tr>
<td>-2917.3</td>
<td>n-9</td>
<td>232000</td>
<td>0.9</td>
<td>5’-rArGrUrArU.............................................-3’</td>
</tr>
<tr>
<td>-2608.5</td>
<td>?</td>
<td>202000</td>
<td>0.8</td>
<td>?</td>
</tr>
<tr>
<td>-2244.6</td>
<td>?</td>
<td>238000</td>
<td>1.0</td>
<td>?</td>
</tr>
<tr>
<td>-1898.1</td>
<td>n-6</td>
<td>368000</td>
<td>1.5</td>
<td>5’-rArGrGrArGrUrArU.....................................-3’</td>
</tr>
<tr>
<td>-1553.4</td>
<td>n-5</td>
<td>316000</td>
<td>1.3</td>
<td>5’-rGrArGrGrArGrUrArU.....................................-3’</td>
</tr>
<tr>
<td>-1246.5</td>
<td>n-4</td>
<td>268000</td>
<td>1.1</td>
<td>5’-rUrGrArGrGrArGrUrArU.....................................-3’</td>
</tr>
<tr>
<td>-918.2</td>
<td>n-3</td>
<td>407000</td>
<td>1.6</td>
<td>5’-rArUrGrArGrGrArGrUrArU.....................................-3’</td>
</tr>
<tr>
<td>-611.7</td>
<td>n-2</td>
<td>382000</td>
<td>1.5</td>
<td>5’-rUrArUrGrArGrGrArGrUrArU.................................-3’</td>
</tr>
<tr>
<td>-345.9</td>
<td>n-rG (internal)</td>
<td>347000</td>
<td>1.4</td>
<td>?</td>
</tr>
<tr>
<td>-329.6</td>
<td>n-rA (internal)</td>
<td>302000</td>
<td>1.2</td>
<td>?</td>
</tr>
<tr>
<td>-306.9</td>
<td>n-1</td>
<td>1180000</td>
<td>4.8</td>
<td>5’-rCrUrArGrArGrGrArGrUrArU.................................-3’</td>
</tr>
<tr>
<td>0</td>
<td>FLP</td>
<td>183000000</td>
<td>74.0</td>
<td>5’-rCrUrArGrArGrGrArGrUrArU.................................-3’</td>
</tr>
<tr>
<td>17.9</td>
<td>+ H2O</td>
<td>1120000</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>304.7</td>
<td>n+rC</td>
<td>160000</td>
<td>0.6</td>
<td>?</td>
</tr>
<tr>
<td>345.5</td>
<td>n+rG</td>
<td>170000</td>
<td>0.7</td>
<td>?</td>
</tr>
<tr>
<td>609.6</td>
<td>n+2</td>
<td>127000</td>
<td>0.5</td>
<td>?</td>
</tr>
</tbody>
</table>
Cas12a Guide RNA Purest Fraction LC/MS Analysis

Average mass spectrum

Deconvolution
Cas12a Guide RNA Purest Fraction Deconvoluted Mass Spectrum

- **FLP**
- **Metal adducts**
- **In-source depurination**
- **Depurination + depyrimidination**
- **2x depyrimidination**
- **n-rUrA**
- **n-rU**
- **n-rG**
- **N+rU**
- **N+rG**

Mass (Da):
- 12000
- 12250
- 12500
- 12750
- 13000

%:
- 0
- 100
# Guide RNA Purest Fraction: Cas12a

<table>
<thead>
<tr>
<th>Mass Delta from FLP (Da)</th>
<th>Identity</th>
<th>Peak Height</th>
<th>%Peak Height</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>-635.1</td>
<td>n-2 (-rUrA)</td>
<td>9.85E+05</td>
<td>0.5</td>
<td>5’-rA.............................................................................................-3’</td>
</tr>
<tr>
<td>-345.6</td>
<td>n-rg (internal)</td>
<td>8.78E+05</td>
<td>0.5</td>
<td>?</td>
</tr>
<tr>
<td>-306.3</td>
<td>n-1 (-rU)</td>
<td>4.04E+06</td>
<td>2.2</td>
<td>5’-rArA.............................................................................................-3’</td>
</tr>
<tr>
<td>-224.1</td>
<td>2x depyrimidination</td>
<td>9.94E+05</td>
<td>0.5</td>
<td>?</td>
</tr>
<tr>
<td>-111.9</td>
<td>depyrimidination</td>
<td>6.41E+06</td>
<td>3.6</td>
<td>?</td>
</tr>
<tr>
<td>0</td>
<td>FLP</td>
<td>1.66E+08</td>
<td>92.3</td>
<td>5’-UrArA.............................................................................................-3’</td>
</tr>
<tr>
<td>+305.7</td>
<td>n+rU</td>
<td>3.99E+05</td>
<td>0.2</td>
<td>?</td>
</tr>
<tr>
<td>+346.2</td>
<td>N+rG</td>
<td>2.88E+05</td>
<td>0.2</td>
<td>?</td>
</tr>
</tbody>
</table>

Higher purity
Impurities on the 5’ end -> impact on activity, not specificity
Improved Safety Profile due to Reduced Potential Off-Target Editing in Human Genome Caused by Errant Guide

**Cas12a Shorter gRNA: Higher Purity and Superior Fidelity**

- **Cas12a**
  - Shorter 41mer gRNA (higher purity & yield)
  - Significantly fewer low-level impurities
    - Final Purity >80% achievable
  - Targeting sequence on the 3'-terminus of molecule
  - 3'-terminus: beginning of synthesis → highest fidelity
  - Impurities likely to impact binding (activity) rather than result in off-target cuts

- **Cas9**
  - Longer 100mer gRNA (lower purity & yield)
  - More impurities, harder to separate chromatographically
    - Final Purity >80% unlikely
  - Targeting sequence on the 5'-terminus of molecule
  - 5' terminus: end of synthesis → lowest fidelity
  - Possible problematic impurities → higher risk of catastrophic off-targets
  - Possible solutions:
    - ligation of shorter RNAs (chemical or enzymatic)
    - In vitro transcription

---

Acknowledgments

Pranjali Ghude
Wesley Mayfield
Steve Wolk
Chrysa Latrick
Bryant Chica
Alexis Benedis
McKenzie Weiss
Shelby Beer
Mark Jones
Stephen Pietrasiewicz
John Zuris
Bruce Eaton
Thank you!

Questions?