Supporting Information for:

Towards the Dual Recursive Deconvolution (DRED) of Resin-Supported Combinatorial Libraries: A Non-Invasive Methodology Based on Bead Self-Encoding and Multispectral Imaging

Hicham Fenniri,* Hartmut G. Hedderich, Kenneth S. Haber, Jihane Achkar, Brian Taylor, Dor Ben-Amotz
Figure S1. Illustration of the dual recursive deconvolution (DRED) strategy. Three spectroscopically distinguishable resin beads (black, white and yellow spheres) are used to encode the first building blocks (A, F, L) of the synthesis. X denotes any of the three building blocks A, F or L. The last building block of an active member of the library could be revealed by pool assay after the last step of the split synthesis, while the first building block could be unveiled through multispectral imaging of the active beads. The green boxes highlight the building blocks required for activity.

Figure S1 outlines the key features of the DRED of a hypothetical 27-member combinatorial library generated through split synthesis. The last step of this process would generate 3 pools each containing 9 compounds.
Screening of each of the pools separately would identify the best third position. The key feature of the DRED method lies in the chemical nature of the beads used. In this case each bead encodes and thus identifies the first randomized building block. In the example of Figure S1, three spectroscopically distinguishable resin beads encode building blocks A, F and L at position 1. As a result the DRED of a library would operate through the identification of the last (pool assay) and first (encoded beads) randomized positions. This process could then be repeated iteratively for the remaining unidentified positions until the entire sequence of the active library-member(s) is unveiled.

Table 1 summarizes the synthetic effort required for the DRED of combinatorial libraries varying in number of steps from 3 to 6 and using 10 or 20 building blocks. The number of steps for the preparation of a library using the split synthesis (column 4) varies linearly while the size of the library increases exponentially (column 3). The number of compounds synthesized per chemical step (column 5) increases rapidly as the library size increases, thereby highlighting the strength of the split synthesis method. Likewise, columns 6 and 7 show a similar trend except that in this case the ratio of compounds synthesized to the number of chemical steps includes the DRED synthetic steps and hence the full identification of the active member of the library.
Table 1. Synthetic effort for dual recursive deconvolution (DRED) of combinatorial libraries.

<table>
<thead>
<tr>
<th>L</th>
<th>N</th>
<th>M</th>
<th>S₁</th>
<th>R₁</th>
<th>S₂</th>
<th>R₂</th>
<th>Libraries to be synthesized</th>
</tr>
</thead>
</table>
| 3 | 10 | 1x10³ | 30 | 33 | 42 | 24 | A₁-X-A₃
   |   |     |    |    | (S₁+12) |    | A₁-A₂-A₃ |
|   | 20 | 8x10³ | 60 | 133 | 82 | 98 | A₁-X-A₃
   |   |     |    |    | (S₁+22) |    | A₁-A₂-A₃ |
| 4 | 10 | 1x10⁴ | 40 | 250 | 62 | 161 | A₁-X-X-A₄
   |   |     |    |    | (S₁+22) |    | A₁₋A₂₋A₃₋A₄ |
|   | 20 | 1.6x10⁵ | 80 | 2,000 | 122 | 1,312 | A₁-X-X-A₄
   |   |     |    |    | (S₁+42) |    | A₁₋A₂₋A₃₋A₄ |
| 5 | 10 | 1x10⁵ | 50 | 2,000 | 96 | 1,042 | A₁-X-X-X-A₅
   |   |     |    |    | (S₁+32+14) |    | A₁₋A₂₋X₋A₃₋A₅ |
|   | 20 | 3.2x10⁶ | 100 | 32,000 | 186 | 17,204 | A₁-X-X-X-A₅
   |   |     |    |    | (S₁+62+24) |    | A₁₋A₂₋X₋A₃₋A₅ |
| 6 | 10 | 1x10⁶ | 60 | 16,667 | 126 | 7,937 | A₁-X-X-X-X-A₆
   |   |     |    |    | (S₁+42+24) |    | A₁₋A₂₋X₋X₋A₃₋A₆ |
|   | 20 | 64x10⁶ | 120 | 533,333 | 246 | 260,163 | A₁₋A₂₋X₋X₋X₋A₄₋A₆
   |   |     |    |    | (S₁+82+44) |    | A₁₋A₂₋X₋X₋X₋A₃₋A₆₋A₄₋A₆ |

[a] L is the number of synthetic steps per library member. [b] N is the number of building blocks. [c] M = N^L is the library size using split synthesis. [d] S₁ = N·L is the number of synthetic steps for the preparation of the library using the split synthesis method. [e] R₁ = M·(S₁)^-1 corresponds to the number of compounds generated per synthetic step. [f] S₂ is the total number of synthetic steps for library and sub-library synthesis, and for full identification of an active member of the library using DRED. [g] R₂ = M·(S₂)^-1 is the number of compounds synthesized and screened per synthetic step. As indicated by this ratio the larger the library the fewer the number of steps for its preparation and for the identification of the active members. [h] General formula of the libraries and sub-libraries to be synthesized using split synthesis and DRED for complete chemical identification of an active member of the library. X denotes randomized positions to be identified in subsequent sub-libraries. Aₙ denotes randomized positions to be identified at that step. Bold Aₙ denotes positions identified in preceding sub-library steps.

For instance, the synthesis and screening of a 64-million member library would barely double the number of chemical steps required for the synthesis of the library using the split synthesis method (246 versus 120), which in
The last column of Table 1 shows the general formula of the libraries and sub-libraries to be synthesized with three to six building blocks or chemical transformations per member using split synthesis and DRED.

The NIRIM instrument uses near infrared (NIR) external cavity narrow band, 400 mW, 785 nm diode laser (SDL-8630), which maximizes resolution and reduces sample fluorescence interference. The charge coupled device (CCD) detector (Princeton instruments LN/CCD-1024 EHRB) has a deep depletion, back illuminated chip which is NIR anti-reflection coated and roughened to virtually eliminate etaloning artifacts (quantum efficiency of 85% at 785 nm and 20% at 1050 nm). The NIRIM also uses a Kaiser Holoscop Imaging spectrograph with an input lens focal length of 75 mm and f/1.4, and an output lens focal length of 85 mm and f/1.4. The image quality of this spectrograph is sufficient to image each 50 μm diameter fiber on a 2x2 pixel (about 54x54 μm) region of the CCD. Note that because the input and output focal lengths are not the same, the spectrograph has a magnification of 1.13, which restricts the number of FIC fibers that may be simultaneously detected to about 80 (representing a rectangular 8x10 fiber region at the collection end of the FIC fiber bundle). Larger spectral images are obtained simply by raster-scanning the sample over an array of adjacent rectangular regions, and concatenating the resulting single-frame images to form a spectral image of an arbitrarily large area. An Nxn image is assembled from Nxn×80 pixels; each
pixel is in fact a 900 channel wide Raman spectrum, the Raman shifts window is from 100 cm\(^{-1}\) to 1900 cm\(^{-1}\). A review of all the remaining components of the NIRIM instrument, including mirrors, lenses, holographic filter, excitation fiber set up and other design considerations were reported recently.\(^{[1]}\)

The resin supported amino-acids studied (Advanced ChemTech) are: Boc-Ala-O-Merrifield (0.9 mmolg\(^{-1}\)); Boc-Asn-O-Merrifield (0.6 mmolg\(^{-1}\)); Boc-Asp(OBzl)-O-Merrifield (1 mmolg\(^{-1}\)); Boc-Cys(Acm)-O-Merrifield (0.8 mmolg\(^{-1}\)); Boc-Gln-O-Merrifield (0.6 mmolg\(^{-1}\)); Boc-Glu(OBzl)-O-Merrifield (0.8 mmolg\(^{-1}\)); Boc-His(DNP)-O-Merrifield (0.6 mmolg\(^{-1}\)); Boc-Ile-O-Merrifield (0.9 mmolg\(^{-1}\)); Boc-Leu-O-Merrifield (1 mmolg\(^{-1}\)); Boc-Lys(2-Cl-Z)-O-Merrifield (0.5 mmolg\(^{-1}\)); Boc-Met-O-Merrifield (0.9 mmolg\(^{-1}\)); Boc-Phe-O-Merrifield (0.8 mmolg\(^{-1}\)); Boc-Pro-O-Merrifield (0.9 mmolg\(^{-1}\)); Boc-Ser(OBzl)-O-Merrifield (0.6 mmolg\(^{-1}\)); Boc-Thr(OBzl)-O-Merrifield (0.6 mmolg\(^{-1}\)); Boc-Trp(O-Merrifield (0.6 mmolg\(^{-1}\)); Boc-Tyr(2-Br-Z)-O-Merrifield (0.6 mmolg\(^{-1}\)); Boc-Val-O-Merrifield (0.8 mmolg\(^{-1}\)). The unsubstituted resins studied are: TentaGel-S-OH (130 \(\mu\)m), 0.3 mmolg\(^{-1}\); and Hydroxymethyl-polystyrene (~90 \(\mu\)m), 1.1 mmolg\(^{-1}\), 1% cross-linked.

The following resins were used in the imaging experiments of Figure 2: (a) 4-Bromo-PS 200-400 mesh (2.5 mmolg\(^{-1}\), Chem-Impex International); (b) 4-Carboxy-PS 100-200 mesh (3.5 mmolg\(^{-1}\), Novabiochem); (c) PEG cross-linked PS 100-200 mesh (2 mmolg\(^{-1}\), Advanced ChemTech); (d) Amino-PEGA (0.4 mmolg\(^{-1}\), Novabiochem); (e) HMBA-SPAR 50 100-200 mesh (polyacrylamide resin, 0.3 mmolg\(^{-1}\), Advanced ChemTech); (f) SPAR 50 200-400 mesh (0.8 mmolg\(^{-1}\), Advanced ChemTech).
The bead samples were placed on a sapphire single crystal\cite{2} positioned in the field of view of the NIRIM and images were recorded. The software used to either acquire or process the experimental data on the NIRIM instrument are pls_image.vi (data acquisition),\cite{3} nirim.vi (3-D data cube acquisition)\cite{3} and MultiSpec\cite{4} (spectral imaging analysis and classification). The latter program requires the user to first select known regions of the image and identify their composition (training fields). The program then uses built-in algorithms (operator's choice) to statistically determine the most likely chemical identity for each fiber's Raman output in the image. The image is then redisplayed with the fibers' Raman output color-coded as to their most likely chemical identity. In this study the images were analyzed using the spectral angle mapping (SAM) algorithm, and the training fields were those of authentic samples of the beads.

References

**Figure 2S.** White light image of 20 beads composed of 6 different resins (same as Figure 2I, main text). The numbers on each bead refer to the single bead near IR-Raman spectra listed below.
Bead#1: 4-Bromo-PS
10X Objective
60s exposure time

Wavenumbers/cm⁻¹

Bead#2: PEG crosslinked-PS
20X objective
60s exposure time

Wavenumbers/cm⁻¹

Bead#3: Amino-PEGA
20X objective
60s exposure time

Wavenumbers/cm⁻¹
Bead#4: HMBA-SPAR 50
10X objective
60s exposure time

Intensity/counts
Wavenumbers/cm⁻¹

Bead#5: Amino-PEGA
20X objective
60s exposure time

Intensity/counts
Wavenumbers/cm⁻¹

Bead#06: SPAR 50
40X objective
60s exposure time

Intensity/counts
Wavenumbers/cm⁻¹
Bead#07:
SPAR 50
20X objective
60s exposure time

Bead#08:
SPAR 50
20X objective
60s exposure time

Bead#09:
Carboxy-PS
20X objective
60s exposure time

637 cm$^{-1}$
Bead#10: Carboxy-PS
20X objective
60s exposure time

Bead#11: HMBA-SPAR 50
20X objective
60s exposure time

Bead#12: Amino-PEGA
20X objective
60s exposure time
Bead#13: Carboxy-PS
20X objective
60s exposure time

637 cm\(^{-1}\)

Bead#14: HMBA-SPAR 50
20X objective
60s exposure time

854 cm\(^{-1}\)

Bead#15: 4-Bromo-PS
20X objective
60s exposure time

1073 cm\(^{-1}\)
Bead#16: Amino-PEGA
20X objective
60s exposure time
1243 cm\(^{-1}\)
1286 cm\(^{-1}\)

Bead#17: PEG crosslinked-PS
20X objective
60s exposure time
703 cm\(^{-1}\)

Bead#18: HMBA-SPAR 50
20X objective
60s exposure time
854 cm\(^{-1}\)
Bead#19: Carboxy-PS
20X objective
60s exposure time

Bead#20: PEG crosslinked-PS
20X objective
60s exposure time