Protein Arrays as Tools for Signal Transduction and Small Molecule Discovery

Mark T. Bedford
The Bedford Lab works on:

Arginine Methylation

Arginine  MMA  ADMA  SDMA

Signal Transduction
Posttranslational Modifications

Herhaus & Dikic. EMBO R. - 2015
Signal Integration on Chromatin

RTK
RAS
AKT
PTEN

Oncohistones  EZH2  IDH1/2
The Talk

• How Protein Domain Microarrays work

• Using arrays to discover “readers” of different signals

• Using arrays to discover small molecule inhibitors
Protein Domains

Modified Peptides
- pY
- pY
- pT
- pS/T
- pS/T
- pS
- meK
- meK
- meK
- meK
- meK
- meK
- meK
- meK
- meK
- acK
- acK
- UbK
- UbK
- UbK
- UbK
- UbK
- UbK
- UbK
- Bromo
- Tudor
- Chromo
- MBT
- PHD
- BHA
- ANK
- PWWP
- WD40

Peptide Motifs
- NPYX
- PXXP
- PPXY
- PPLP
- PPPR
- FPPPP
- FPXXF
- FPPH
- PTB
- SH3
- WW
- EVH1
- VHS
- GYF
- PDZ

Nucleic Acids and Phosphoinositides
- RNA
- meRNA
- DNA
- meDNA
- PIP
- PIP
- PIP
- PIP
- PIP
- RNA
- KH
- YTH
- Znf
- MBD
- PH
- PX
- BAR
- FYVE
- ENTH
Protein Array & Analysis Core (PAAC)

• We generate libraries of recombinant protein domains.

• These focused arrays include:
  – Yp reading array – SH2 & PTBs (89 domains).
  – Sp/Tp reading array – BRCT, FHA, 14-3-3, Polo box & WW (90).
  – K/Rme reading array – Tudor, Chromo, MBT, YTH, HORMA, & PHD (232).
  – Kac reading array – Bromo & YEATS (53).
  – Proline-rich reading array – WW and SH3 (158).
  – C-terminal reading array – PDZ (98).
  – Ubiquitin-binding domains – UBA, UIM UBX, CUE, GAT & VHS (140)
  – Ordered - RNA binding domains (334) and BAH (40).
  – In pipeline - phospho-lipid binding domains.

Total – 1,234 domains
Some of the Details

- The protein domains are cloned as GST fusions.
- We use Biomatik (Canada) for our gene synthesis projects.
- We clone human domains.
- They are codon optimized for bacterial expression.

- The recombinant domains are arrayed onto nitrocellulose coated glass slides with an Aushon 2470 (pin) microarrayer.

- Arrays are probed with fluorescently labeled peptides or compounds.
Probing the Protein Domain Microarrays
The Talk

• How Protein Domain Microarrays work

• Using arrays to discover “readers” of different signals
  • Phosphorylation
  • Methylation
  • Ubiquitination

• Using arrays to discover small molecule inhibitors
Examples of phospho-dependent interactions
The Histone Code
# Tudor & Chromo Array

## Biomatik Full Array 2

<table>
<thead>
<tr>
<th>TUDOR</th>
<th>CHROMO</th>
<th>AGNET</th>
<th>IDCL</th>
<th>BROMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 53BP1(1-2)</td>
<td>H1 ARID4A</td>
<td>L1 FMR1</td>
<td>M1 PCNA</td>
<td>N1 SPF140*</td>
</tr>
<tr>
<td>A2 53BP1(1-2)*</td>
<td>H2 ARID4B</td>
<td>L2 FXR1</td>
<td>MBT</td>
<td>N2 BRD4(1)*</td>
</tr>
<tr>
<td>A3 TDRD1-2</td>
<td>H3 CBX2</td>
<td>L3 FXR2</td>
<td>M2 L3MBTL1(1-3)</td>
<td>N3 WDR9(1-2)*</td>
</tr>
<tr>
<td>A4 TDRD1-3</td>
<td>H4 CBX4</td>
<td>L4 GLP</td>
<td>M3 L3MBTL1(1-3)*</td>
<td>N4 SMAP*</td>
</tr>
<tr>
<td>A5 TDRD1-4</td>
<td>H5 CBX6</td>
<td>L5 ORC1</td>
<td>M4 L3MBTL3</td>
<td>A1</td>
</tr>
<tr>
<td>A6 TDRD2</td>
<td>H6 CBX7</td>
<td>L6 N-CAPD3(1-4)</td>
<td>M5 DNMT3A*</td>
<td>TDRD9</td>
</tr>
<tr>
<td>A7 TDRD3</td>
<td>H7 CBX8</td>
<td>L7 N-CAPD3(5-8)</td>
<td>M6 NSD1 PWWP+PHD*</td>
<td>PHD</td>
</tr>
<tr>
<td>B1 TDRD4-1</td>
<td>J1 CHD1</td>
<td>L8 N-CAPD4</td>
<td>M7 ZCWPW1</td>
<td>G1 ING2</td>
</tr>
<tr>
<td>B2 TDRD4-3</td>
<td>J2 CHD2</td>
<td>L9 N-CAPD5</td>
<td></td>
<td>G2 TAF3</td>
</tr>
<tr>
<td>B3 TDRD4-4</td>
<td>J3 CHD6</td>
<td>L10 N-CAPD6</td>
<td></td>
<td>G3 TRIM24</td>
</tr>
<tr>
<td>B4 TDRD5</td>
<td>J4 CHD7</td>
<td>L11 N-CAPD7</td>
<td></td>
<td>G4 CHD5(1-2)*</td>
</tr>
<tr>
<td>B5 TDRD6-5</td>
<td>J5 CHD8</td>
<td>L12 N-CAPD8</td>
<td></td>
<td>G5 CHD</td>
</tr>
<tr>
<td>B6 TDRD6-6</td>
<td>J6 CHD9</td>
<td>L13 N-CAPD9</td>
<td></td>
<td>G6 CHD4</td>
</tr>
<tr>
<td>B7 TDRD6-7</td>
<td>J7 TIP60</td>
<td>L14 N-CAPD10</td>
<td></td>
<td>G7 CHD5</td>
</tr>
<tr>
<td>C1 TDRD8</td>
<td>K1 MPP8</td>
<td>L15 N-CAPD11</td>
<td></td>
<td>H1</td>
</tr>
<tr>
<td>C2 TDRD9</td>
<td>K2 MRG15</td>
<td>L16 N-CAPD12</td>
<td></td>
<td>H2</td>
</tr>
<tr>
<td>C3 TDRD10</td>
<td>K3 MYST1/MOF</td>
<td>L17 N-CAPD13</td>
<td></td>
<td>H3</td>
</tr>
<tr>
<td>C4 TDRD11</td>
<td>K4 SMARCC1</td>
<td>L18 N-CAPD14</td>
<td></td>
<td>H4</td>
</tr>
<tr>
<td>C5 TDRD12-1</td>
<td>K5 SMARCC2</td>
<td>L19 N-CAPD15</td>
<td></td>
<td>H5</td>
</tr>
<tr>
<td>C6 ARID4A</td>
<td>K6 SUV39H1</td>
<td>L20 N-CAPD16</td>
<td></td>
<td>H6</td>
</tr>
<tr>
<td>C7 ARID4B</td>
<td>K7 SUV39H2</td>
<td></td>
<td></td>
<td>H7</td>
</tr>
<tr>
<td>D1 AKAP1</td>
<td>L1 FMR1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2 LBR</td>
<td>L2 FXR1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3 PHF1</td>
<td>L3 FXR2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D4 PHF19</td>
<td>L4 GLP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D5 PHF20</td>
<td>L5 ORC1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6 PHF20-2</td>
<td>L6 N-CAPD3(1-4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D7 PHF20L1</td>
<td>L7 N-CAPD3(5-8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1 SETDB1</td>
<td>L8 N-CAPD4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2 SMN</td>
<td>L9 N-CAPD5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3 SMN*</td>
<td>L10 N-CAPD6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4 SPF30</td>
<td>L11 N-CAPD7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E5 UHRF1</td>
<td>L12 N-CAPD8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E6 ZGPAT</td>
<td>L13 N-CAPD9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E7 JMJD2A(1-2)</td>
<td>L14 N-CAPD11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1 JMJD2B</td>
<td>L15 N-CAPD12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2 JMJD2C</td>
<td>L16 N-CAPD13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3 MTF2</td>
<td>L17 N-CAPD14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4 PHF1*</td>
<td>L18 N-CAPD15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5 Pombe1*</td>
<td>L19 N-CAPD16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F6 Spindlin1*</td>
<td>L20 N-CAPD17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1 CHD5</td>
<td>L21 N-CAPD18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2 TAF3</td>
<td>L22 N-CAPD19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3 TRIM24</td>
<td>L23 N-CAPD20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G4 CHD5(1-2)*</td>
<td>L24 N-CAPD21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>ARID4A</td>
<td>L25 N-CAPD22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>ARID4B</td>
<td>L26 N-CAPD23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H3</td>
<td>CBX2</td>
<td>L27 N-CAPD24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H4</td>
<td>CBX4</td>
<td>L28 N-CAPD25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H5</td>
<td>CBX6</td>
<td>L29 N-CAPD26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H6</td>
<td>CBX7</td>
<td>L30 N-CAPD27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H7</td>
<td>CBX8</td>
<td>L31 N-CAPD28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- * = Bedford Lab Construct
- † = Michelle Barton
- ‡ = Or Gozani

- TUDOR = Tudor domain-containing proteins
- CHROMO = Chromo domain-containing proteins
- AGENET = AGENET domain-containing proteins
- IDCL = IDCL domain-containing proteins
- BROMO = BROMO domain-containing proteins
- PHD = Plant homeodomain proteins
- MBT = Multi-branched proteins
- PWWP = Polycomb-related domain proteins
- ZF-CW = ZF-CW domain-containing proteins
- MBT = Multi-branched proteins
Reading the histone code

Kim et. al., EMBO Rep, 2006
Focus on PHDs

H3K4me0

H3K4me3
Ubiquitin-binding array

Ubiquitin-Binding Domains Array

UBA
UIM
GAT
CUE
JAB
VHS
MIU
PFU
UBX
UBM
UBZ
UEV

140 domains
Focus on UBDs

1. ASCC2
2. Cezanne1
3. Cezanne2
4. TDRD3
5. RAP80
6. TOM1L2
7. RNF168

Collaboration with LifeSensors & Bin Wang
Get the Hit – Build the Story
The Talk

• How Protein Domain Microarrays work

• Using arrays to discover “readers” of different signals

• Using arrays to discover small molecule inhibitors
Methyl Readers
Small molecule binding to methyl readers

Collaboration with Stephen Frye
Nature Chemical Biology 2013
Collaboration with Gianluca Sbardella
University of Salerno

the EML series
Loss of specificity

Selective for PHF20
EML405 engages two Tudor domains

Collaboration with Haitao Li
The EML series competition with peptide ligand
RNAseq identified new SPIN1 target genes

Down-regulated genes

<table>
<thead>
<tr>
<th>SPIN1_KD</th>
<th>EML631</th>
</tr>
</thead>
<tbody>
<tr>
<td>936</td>
<td>933</td>
</tr>
<tr>
<td>655</td>
<td></td>
</tr>
</tbody>
</table>

88 expected by chance, 655 observed, 655/88=7.39 fold enrichment, Fisher’s exact test P-value < 2.2e-16
Down-regulated SPIN1 target genes

Bae et. al., Nature Chemical Biology 2017
Protein Domain Microarrays can:

• Discover new Signaling Paradigms

• Discover Small Molecules to block Signaling Pathways
The Players:

Cari Sagum

Jianji Chen

Funding Support:
Center for Cancer Epigenetics
CPRIT
NIH