Predicting the specificity of motif-mediated interactions
Many proteins recognize short linear motifs in the sequences of their polymer partners

PxxP

PPxY

P(T/S)x(L/V/I)P
Structure-based prediction of interaction specificity

• General problem: given a structural model of a protein-{peptide,DNA,RNA} complex, calculate relative affinities for alternate sequences of the partner

• Together with homology modeling, could be used to predict approximate binding specificities for uncharacterized transcription factors, peptide-binding domains, etc, in high-throughput fashion.
Protein-DNA interaction specificity

conformational sampling + DNA sequence optimization
protein-DNA methods

• DNA flexibility
  – wriggling (for small torsion angle changes)
  – double-fragment closure
  – DNA rotamer library
  – double-helical fragments
  – DNA “design” preserving WC base pairing
  – base-centric foldtree

• centroid env and pair for protein-DNA

• waters
Base-centric kinematics

Rigid-body “jump” (6 DOFs)
go to centroid-level fragment simulation
lowest energy model w/ native
waters in pymol
Kinase Specificity

- Globular Domain Motifs (SH2, SH3, PDZ…)
- Linear Docking Motifs (…XXRRXSLXX…)
- Most Thr/Ser kinases use Linear Docking Motifs

Test System

- PKA crystal structure (2.2 Å) in complex with ATP and inhibitor peptide
- Superposition with PKA in complex with ANP-PNP and substrate peptide allows placement of phospho-acceptor in catalytic geometry
Step 1.
Generate Backbone
Step 2.
Refine
Backbone
Step 3. Design Sequence

IN  OUT

Image of molecular structure
Step 4.

Sequence Motif

Predicted Peptides | Score
---|---
RKFSIVH | 7
GRKSPPG | 14
RGSSPPI | 3
HILSGED | 21
SGRSKRI | 9
1. Generate Backbone

- Start with known phospho-acceptor conformation
- Generate ghost peptides via fragment insertion
- Scoring function penalizes clashes and backbone solvent accessibility
2. Refine Backbone

- Assign random sequence and turn on sequence-specific scoring terms
- Optimize backbone structure with small torsional rotations
3. Design Sequence

- Switch to all-atom representation and design optimal sequence
- Refine backbone structure with tiny backbone moves and gradient minimization
- Scoring function uses all-atom terms with soft-core VdW potential
4. Sequence Profile

<table>
<thead>
<tr>
<th>Predicted Peptides</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>RKFSIVH</td>
<td>7</td>
</tr>
<tr>
<td>GRKSPPG</td>
<td>14</td>
</tr>
<tr>
<td>RGSSPPI</td>
<td>3</td>
</tr>
<tr>
<td>HILSGED</td>
<td>21</td>
</tr>
<tr>
<td>SGRSKRI</td>
<td>9</td>
</tr>
</tbody>
</table>
Experimental Verification

- Peptide Microarray

- Peptide Microarray
  - Motif R[RK]X[ST][ILVFY][DCX]XD
  - In vivo kinase target database
    - http://phospho.elm.eu.org/
- ROSETTA prediction
- Peptide Microarray
- *In vivo* kinase target database