Human intervention for FM targets in CASP11

BAKER = BAKER-ROSETTASERVER + Human

HUMANS INCLUDE:
- Sergey Ovchinnikov
- Ray Wang
- Yuan Liu
- Hahnbeom Park
- David Kim
Mainly done by a stateless one

Sergey Ovchinnikov
GREMLIN used to measure Co-evolution

Global statistical model:

\[ P(x_1, x_2) = \frac{1}{Z} \exp \left( \sum_i v_i(x_i) + \sum_{i,j} w_{ij}(x_{ij}) \right) \]

- \( v_i \) = one-body energy (PSSM)
- \( w_{ij} \) = two-body energy (Coupling term)

Kamisetty et al. 2013
Accuracy Depending on Number of Sequences

$L = \text{length of protein}$
Restraints in Rosetta

**SIGMOID**: with no gradient beyond a certain point will require large amount of sampling, but will prevent false positives from affecting results. The correct model will maximize the self-consistent contacts.
Minimize energy w/ Sigmoid Restraint

Starting conformation

Minimize energy w/ Harmonic/Bounded Restraint

Will maximize self-consistent contacts
Amino acid pair specific $\text{C} \beta - \text{C} \beta$ distance cutoffs were used.

<table>
<thead>
<tr>
<th></th>
<th>G</th>
<th>A</th>
<th>S</th>
<th>V</th>
<th>C</th>
<th>T</th>
<th>P</th>
<th>D</th>
<th>N</th>
<th>I</th>
<th>L</th>
<th>Q</th>
<th>M</th>
<th>H</th>
<th>K</th>
<th>F</th>
<th>Y</th>
<th>R</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>4.5</td>
<td>5.2</td>
<td>5.5</td>
<td>5.7</td>
<td>5.8</td>
<td>5.6</td>
<td>6.1</td>
<td>6.1</td>
<td>6.3</td>
<td>6.4</td>
<td>6.6</td>
<td>7.0</td>
<td>7.3</td>
<td>7.4</td>
<td>7.5</td>
<td>8.2</td>
<td>8.0</td>
<td>9.1</td>
<td>9.2</td>
</tr>
<tr>
<td>A</td>
<td>5.2</td>
<td>5.4</td>
<td>5.8</td>
<td>6.1</td>
<td>6.0</td>
<td>6.4</td>
<td>6.4</td>
<td>6.8</td>
<td>6.6</td>
<td>6.7</td>
<td>7.1</td>
<td>7.6</td>
<td>7.6</td>
<td>7.6</td>
<td>8.3</td>
<td>8.2</td>
<td>9.1</td>
<td>9.4</td>
<td>9.3</td>
</tr>
<tr>
<td>S</td>
<td>5.5</td>
<td>5.8</td>
<td>6.2</td>
<td>6.6</td>
<td>6.6</td>
<td>6.5</td>
<td>6.9</td>
<td>6.8</td>
<td>7.1</td>
<td>7.1</td>
<td>7.4</td>
<td>7.5</td>
<td>7.8</td>
<td>8.0</td>
<td>8.1</td>
<td>8.8</td>
<td>8.7</td>
<td>9.6</td>
<td>9.8</td>
</tr>
<tr>
<td>V</td>
<td>5.7</td>
<td>5.9</td>
<td>6.6</td>
<td>6.8</td>
<td>6.9</td>
<td>6.8</td>
<td>7.1</td>
<td>7.0</td>
<td>7.2</td>
<td>7.4</td>
<td>7.6</td>
<td>7.4</td>
<td>8.0</td>
<td>8.3</td>
<td>8.2</td>
<td>8.1</td>
<td>9.1</td>
<td>9.4</td>
<td>9.5</td>
</tr>
<tr>
<td>C</td>
<td>5.8</td>
<td>6.1</td>
<td>6.6</td>
<td>6.9</td>
<td>6.4</td>
<td>6.8</td>
<td>7.2</td>
<td>7.0</td>
<td>7.2</td>
<td>7.5</td>
<td>7.7</td>
<td>7.4</td>
<td>8.0</td>
<td>8.3</td>
<td>8.4</td>
<td>8.5</td>
<td>9.0</td>
<td>9.4</td>
<td>9.5</td>
</tr>
<tr>
<td>T</td>
<td>5.6</td>
<td>6.0</td>
<td>6.5</td>
<td>6.8</td>
<td>6.8</td>
<td>6.7</td>
<td>7.1</td>
<td>7.0</td>
<td>7.2</td>
<td>7.4</td>
<td>7.6</td>
<td>7.6</td>
<td>8.1</td>
<td>8.4</td>
<td>8.2</td>
<td>8.7</td>
<td>9.0</td>
<td>9.8</td>
<td>9.8</td>
</tr>
<tr>
<td>P</td>
<td>6.1</td>
<td>6.4</td>
<td>6.9</td>
<td>7.1</td>
<td>7.2</td>
<td>7.1</td>
<td>7.3</td>
<td>7.3</td>
<td>7.5</td>
<td>7.6</td>
<td>7.8</td>
<td>7.9</td>
<td>8.3</td>
<td>8.2</td>
<td>8.5</td>
<td>9.2</td>
<td>8.9</td>
<td>10.0</td>
<td>10.3</td>
</tr>
<tr>
<td>D</td>
<td>6.1</td>
<td>6.4</td>
<td>6.8</td>
<td>7.0</td>
<td>7.0</td>
<td>7.3</td>
<td>8.0</td>
<td>7.7</td>
<td>7.5</td>
<td>7.7</td>
<td>8.9</td>
<td>8.6</td>
<td>8.4</td>
<td>8.6</td>
<td>9.3</td>
<td>9.1</td>
<td>10.0</td>
<td>10.1</td>
<td>9.9</td>
</tr>
<tr>
<td>N</td>
<td>6.3</td>
<td>6.8</td>
<td>7.1</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.7</td>
<td>7.7</td>
<td>7.6</td>
<td>7.9</td>
<td>8.5</td>
<td>8.5</td>
<td>8.6</td>
<td>8.7</td>
<td>9.3</td>
<td>9.2</td>
<td>10.0</td>
<td>10.1</td>
<td>10.0</td>
</tr>
<tr>
<td>I</td>
<td>6.4</td>
<td>6.6</td>
<td>7.1</td>
<td>7.4</td>
<td>7.5</td>
<td>7.4</td>
<td>7.6</td>
<td>7.5</td>
<td>7.6</td>
<td>8.1</td>
<td>8.3</td>
<td>7.9</td>
<td>8.3</td>
<td>8.9</td>
<td>8.5</td>
<td>8.3</td>
<td>9.6</td>
<td>9.7</td>
<td>9.7</td>
</tr>
<tr>
<td>L</td>
<td>6.6</td>
<td>6.7</td>
<td>7.4</td>
<td>7.6</td>
<td>7.7</td>
<td>7.6</td>
<td>7.8</td>
<td>7.7</td>
<td>7.9</td>
<td>8.3</td>
<td>8.5</td>
<td>8.1</td>
<td>8.5</td>
<td>9.1</td>
<td>8.7</td>
<td>8.5</td>
<td>9.9</td>
<td>9.9</td>
<td>9.9</td>
</tr>
<tr>
<td>E</td>
<td>7.0</td>
<td>7.1</td>
<td>7.5</td>
<td>7.4</td>
<td>7.4</td>
<td>7.6</td>
<td>7.9</td>
<td>8.9</td>
<td>8.5</td>
<td>7.9</td>
<td>8.1</td>
<td>9.9</td>
<td>9.3</td>
<td>8.9</td>
<td>9.5</td>
<td>9.8</td>
<td>9.4</td>
<td>10.5</td>
<td>10.7</td>
</tr>
<tr>
<td>Q</td>
<td>7.3</td>
<td>7.6</td>
<td>7.8</td>
<td>8.0</td>
<td>8.0</td>
<td>8.1</td>
<td>8.3</td>
<td>8.6</td>
<td>8.5</td>
<td>8.3</td>
<td>8.5</td>
<td>9.3</td>
<td>9.1</td>
<td>9.1</td>
<td>9.4</td>
<td>9.7</td>
<td>9.5</td>
<td>10.5</td>
<td>10.6</td>
</tr>
<tr>
<td>M</td>
<td>7.4</td>
<td>7.6</td>
<td>8.0</td>
<td>8.3</td>
<td>8.3</td>
<td>8.4</td>
<td>8.2</td>
<td>8.4</td>
<td>8.6</td>
<td>8.9</td>
<td>9.1</td>
<td>8.9</td>
<td>9.1</td>
<td>9.5</td>
<td>9.4</td>
<td>9.1</td>
<td>10.3</td>
<td>10.4</td>
<td>10.3</td>
</tr>
<tr>
<td>H</td>
<td>7.5</td>
<td>7.6</td>
<td>8.1</td>
<td>8.2</td>
<td>8.4</td>
<td>8.2</td>
<td>8.5</td>
<td>8.6</td>
<td>8.7</td>
<td>8.5</td>
<td>8.7</td>
<td>9.5</td>
<td>9.4</td>
<td>9.4</td>
<td>10.6</td>
<td>9.6</td>
<td>9.6</td>
<td>10.8</td>
<td>10.9</td>
</tr>
<tr>
<td>K</td>
<td>8.2</td>
<td>8.3</td>
<td>8.8</td>
<td>8.1</td>
<td>8.5</td>
<td>8.7</td>
<td>9.2</td>
<td>9.3</td>
<td>9.3</td>
<td>8.3</td>
<td>8.5</td>
<td>9.8</td>
<td>9.7</td>
<td>9.1</td>
<td>9.6</td>
<td>10.7</td>
<td>9.3</td>
<td>10.6</td>
<td>11.3</td>
</tr>
<tr>
<td>F</td>
<td>8.0</td>
<td>8.2</td>
<td>8.7</td>
<td>9.1</td>
<td>9.0</td>
<td>9.0</td>
<td>8.9</td>
<td>9.1</td>
<td>9.2</td>
<td>9.6</td>
<td>9.9</td>
<td>9.4</td>
<td>9.5</td>
<td>10.3</td>
<td>9.6</td>
<td>9.3</td>
<td>10.9</td>
<td>11.0</td>
<td>10.6</td>
</tr>
<tr>
<td>Y</td>
<td>9.1</td>
<td>9.1</td>
<td>9.6</td>
<td>9.4</td>
<td>9.4</td>
<td>9.8</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>9.7</td>
<td>9.9</td>
<td>10.5</td>
<td>10.5</td>
<td>10.4</td>
<td>10.8</td>
<td>10.6</td>
<td>11.0</td>
<td>11.5</td>
<td>11.6</td>
</tr>
<tr>
<td>R</td>
<td>9.2</td>
<td>9.4</td>
<td>9.8</td>
<td>9.5</td>
<td>9.5</td>
<td>9.8</td>
<td>10.3</td>
<td>10.1</td>
<td>10.1</td>
<td>9.7</td>
<td>9.9</td>
<td>10.7</td>
<td>10.6</td>
<td>10.3</td>
<td>10.9</td>
<td>11.3</td>
<td>10.6</td>
<td>11.6</td>
<td>12.1</td>
</tr>
<tr>
<td>W</td>
<td>9.0</td>
<td>9.3</td>
<td>9.8</td>
<td>10.0</td>
<td>9.8</td>
<td>10.0</td>
<td>9.7</td>
<td>9.9</td>
<td>10.0</td>
<td>10.5</td>
<td>10.7</td>
<td>10.3</td>
<td>10.4</td>
<td>11.1</td>
<td>10.7</td>
<td>10.1</td>
<td>11.8</td>
<td>11.8</td>
<td>11.4</td>
</tr>
</tbody>
</table>

Calculated by determining the maximum distance $\text{C} \beta - \text{C} \beta$ can be at and still have a minimal distance of 5.0Å
HUMAN -> Sergey’s protocol

**TARGET**

HHblits
Jackhmmer

**ALIGNMENT**

Identity filter 90%
Coverage 75%

**CONTACT PREDICTIONS**

GREMLIN

If >= 1L Sequences

**BAKER SERVER MODELS**

If MOST contacts made

**SCORE**

**SUBMIT**

If MOST contacts not made

RosettaAB

**MORE MODELS**

If SOME contacts not made

Iterative RosettaCM to refine

By default we submit the BAKER_SERVER models, unless there is Co-evolution data and HUMAN feels SERVER didn’t sample enough to make all the contacts.

Note, server models were NOT re-ranked.

$L =$ length of protein
RESULTS

GDT_TS Best of 5

GDT_TS Best of 5
T0806 predicted contacts

YAAA_ECOLI
6.4L Sequences
Top 1.5L contacts
RosettaAB + Relax T0806

Native

Top 4 models
None of the models made all non-local contacts

Contact maps of the top 4 models

- Structure Contacts (5Å)
- Predicted Contacts
Pipeline

Abinitio (using RosettaAB)

Hybridize (using RosettaCM)
Fragment insertion (20 trials)

Repeat until deadline or convergence.
T0806 prediction

RMSD: 3.576
GDT_TS: 0.6035

NATIVE
MODEL
Post-CASP repeat until **convergence**

Post-CASP results

**RMSD:** 2.635  
**GDT_TS:** 0.6836
T0806 overlay of predicted contacts on Native
T0824 predicted contacts

NucB
2.5L Sequences
Top 1.5L contacts
T0824 RosettaAB + Relax
Error in model due to bad fragments, Due to incorrect PSIPRED predictions.

RMSD: 4.167
GDT_TS: 0.5532
T0824 overlay of predicted contacts on Native
Acknowledgements

Hetunandan Kamichetty (GREMLIN)

Rosetta Commons
David Baker

Rosetta@home users for generous computing resources

CASP organizers, assessors, structural biologists who provided structures

Andriy Kryshtafovych  <-  I think Sergey had a lot of questions