

# Übungen zur Vorlesung Algorithmische Bioinformatik

Freie Universität Berlin, WS 2010/11  
Roland Krause · Martin Vingron · Matthias Winkelmann

**Blatt 13 · Ausgabe am 25.1.2011**  
**Abgabe am 31.1.2011 vor Beginn der Vorlesung**

**Aufgabe 43** (Homology Modelling). Eine gerne für Homology Modelling verwendete Software ist *Modeller*. Installieren Sie die Software<sup>1</sup>. Sie benötigen zur Installation einen Lizenzschlüssel, den Sie kostenlos erhalten können<sup>2</sup>. Verwenden Sie unbedingt Ihre E-Mailadresse an der FU, da Sie nur so einer akademischen Institution zugeordnet werden können.

Sie sollen nun eine Strukturvorhersage für folgende Sequenz erstellen:

```
MTATAVALATTAVHAADDLKIYKGTGPLEAYAKQTETGLMMGLEAYTKGTMTLDGRK
IVVITKDDQSKPDLKAALAEAYQDDGADIAIGTSSSAAALADLPVAEENKKILIVEPAV
ADQITGEKWNRYIFRTGRNSSQDAISNAVAIGKQGVTIATLAQDYAFGRDGVAAFKEALA
KTGATLATEEYVPTTTTDFTAVGQRLFDALKDKPGKKIIVWVWAGGGDPLTKLQDMDPKR
YGIELSTGGNILPALAAYKRLPGMEGATYYYYDIPKNPINEWLVTEHQKRFNAPPDFFTA
GGFSAAMAVVTAVQKAKSTDTEKLIAAMEGMEFDTPKGMVFRKEDHQALQSMYHFKVKV
DPAVAWAVLEPVRKIEEMNIPIKNKK
```

## Hinweise

- Lesen Sie das Einführungstutorial<sup>3</sup>. Sie können die dort verwendeten Scripte mit geringen Anpassungen wiederverwenden.
- Die Datei `pdb.95.pir` ist veraltet. Verwenden Sie stattdessen eine aktuellere<sup>4</sup>. Im Skript muss das Dateiformat dementsprechend auf FASTA angepasst werden.
- Wählen Sie Ihre beste Vorhersage aus und laden Sie sie sowie die echte Struktur 3NOX in PyMol. Führen Sie dort ein Strukturalignment aus. Welchen RMSD erhalten Sie?
- (Optional) Wie gut hätten Sie in CASP abgeschnitten? Berechnen Sie mit `maxcluster`<sup>5</sup> den GDT\_TS Ihrer Vorhersage und vergleichen Sie ihn mit den Ergebnissen<sup>6</sup>.

**Aufgabe 44** (Validierung). In der vorherigen Aufgabe haben Sie bereits den DOPE-Score benutzt, um die Qualität Ihrer Vorhersage zu bewerten. Nun sollen Sie bewerten, wie gut DOPE zwischen guten und schlechten Vorhersagen diskriminieren kann.

- Laden Sie 30 der Vorhersagen für das CASP Target T0552 herunter<sup>7</sup>. Wählen Sie dabei in gleicher Menge zufällige Modelle mit GDT\_TS-Scores über und unter 50. Vorhersagen mit GDT\_TS über 50 klassifizieren wir als gut. Andere als schlecht.

---

<sup>1</sup>[http://salilab.org/modeller/download\\_installation.html](http://salilab.org/modeller/download_installation.html)

<sup>2</sup><http://salilab.org/modeller/registration.html>

<sup>3</sup><http://salilab.org/modeller/tutorial/basic.html>

<sup>4</sup>[http://lectures.molgen.mpg.de/Algorithmische\\_Bioinformatik\\_WS1011/dokumente/pdb.fasta](http://lectures.molgen.mpg.de/Algorithmische_Bioinformatik_WS1011/dokumente/pdb.fasta)

<sup>5</sup><http://www.sbg.bio.ic.ac.uk/~maxcluster/index.html>

<sup>6</sup>[http://predictioncenter.org/casp9/results.cgi?view=tables&target=T0528&model=1&groups\\_id=](http://predictioncenter.org/casp9/results.cgi?view=tables&target=T0528&model=1&groups_id=)

<sup>7</sup>[http://predictioncenter.org/casp9/results.cgi?view=tables&target=T0552-D1&model=1&groups\\_id=](http://predictioncenter.org/casp9/results.cgi?view=tables&target=T0552-D1&model=1&groups_id=)

- Erstellen Sie ein Modeller-Script, das den DOPE-Score einer Vorhersage berechnet. Sie können sich dabei am Script `evaluate_model.py` aus dem Modeller-Tutorial orientieren, müssen jedoch keine Positionsgenauen Scores berechnen und sollten den normalisierten Score verwenden: `score = mdl.assess_normalized_dope()`.
- Plotten Sie eine ROC-Curve und bewerten Sie die Diskriminanzfähigkeit von DOPE.
- Wählen Sie einen geeigneten *Cutoff*. Stellen Sie die Kontingenztabelle auf und berechnen Sie Sensitivität und Spezifität.
- Schlagen Sie eine andere Methode vor, die Güte von DOPE zu bewerten.
- (Optional) Rücken Sie DOPE in ein (noch?) besseres Licht! Verwenden Sie beliebige Strukturvorhersagen<sup>8</sup> von der CASP-Seite und echte Proteinstrukturen der PDB. Ermitteln Sie die Fähigkeit von DOPE, zwischen diesen zu unterscheiden.

---

<sup>8</sup><http://predictioncenter.org/casp9/results.cgi?offset=T0545&>