

# AST Übung 6 - Homology Modelling

## Einführung

Das Ziel dieser Übung ist es sich mit einem einfachen *Homology Modelling* einer Proteinstruktur mit Hilfe von Chimera und dem *Modeller* Programm vertraut zu machen. Dabei wird für eine kleine bekannte Proteinstruktur mittels BLAST in der PDB nach potentiellen Templates gesucht und mit diesen ein Sequenzalignment durchgeführt. Basierend auf dem Alignment erstellt *Modeller* dann Vorhersagemodelle, welche evaluiert werden.

## Zielsequenz und Template Suche

Laden Sie die Struktur des Ferredoxins 1dur in *Chimera*. Im Folgenden wird davon ausgegangen, dass die Struktur des Proteins nicht bekannt ist, schalten Sie daher dessen Sichtbarkeit aus. Es wird jedoch die Sequenz für die Modellierung und die Struktur später zur Bewertung der Vorhersagen benötigt.

Führen Sie mit Hilfe des Blast Protein Tools eine BLAST-Suche der PDB durch. Das beste Ergebnis wird offensichtlich die offene Struktur sein – ignorieren Sie diese. Laden Sie die Struktur des nächstbesten Ergebnisses 1fca, in dem Sie auf den Load Structure Button klicken (bei 1c1f handelt es sich um eine mit NMR ermittelte Struktur, die wir ebenfalls ignorieren). Diese Struktur wird im Folgenden als Template verwendet. Führen Sie ein Sequenzalignment beider Strukturen durch:

Tools -> Sequence -> Align Chain Sequences

Wählen Sie beide Modelle unter Chains to align aus, behalten Sie die Standardeinstellungen und klicken Sie auf OK. Betrachten Sie das Alignment und schätzen Sie die Ähnlichkeit der Sequenzen ein. Berechnen Sie diese anschließend über Info -> Percent Identity..., wo Sie das erste Modell mit dem Zweiten vergleichen und das Ergebniss durch die kleinere Sequenzlänge teilen. Unten links im Alignmentfenster (MUSCLE Alignment) erscheint die berechnete Sequenzidentität. Reicht diese um 1fca als Template zu verwenden?

## MODELLER

Behalten Sie dieses Alignmentfenster bis zum Ende der Übung offen. Starten Sie von dort aus das *Modeller* Tool:

Structure -> Modeller (homology)

Wählen Sie als Zielsequenz 1dur und als Template 1fca (Modell muss in der Templateliste angeklickt werden) aus. Wählen Sie die Option Run Modeller locally aus und geben Sie für Location of Modeller executable folgenden Pfad an:

```
/home/torda/bin/modeller9.14/bin/mod9.14
```

Schauen Sie sich auch die Optionen unter Advanced Options an, lassen Sie aber die Standardeinstellungen. Klicken Sie auf Apply um eine Modellierung auszuführen.

Die Modellierung kann einige Minuten dauern. Die erstellten Vorhersagemodelle werden in Chimera geladen und es erscheint ein Fenster in dem die Ergebnisse durch zwei Metriken bewertet werden, welche auf statistischem Potential beruhen (Annäherung an Energiefunktionen basierend auf Statistiken die von PDB Strukturen hergeleitet wurden). Ein GA341 Wert von  $> 0.7$  deutet dabei ein zuverlässiges Modell an, während für zDOPE (*normalized Discrete Optimized Protein Energy*) negativere Werte bessere Modelle bedeuten. Beides bewertet dabei nur die Qualität der erstellten Strukturen, nicht die Ähnlichkeit zum Zielmodell (da dieses unbekannt ist). Notieren Sie sich die Ergebnisse oder behalten Sie das Ergebnisfenster bis zum Ende der Übung offen.

## Vergleich mit Vorhersage

Deaktivieren Sie die Sichtbarkeit des Templatemodells und aktivieren Sie das Zielmodell. Überlagern Sie falls notwendig die Vorhersagemodelle mit dem Zielmodell mit Hilfe des MatchMaker Tools und vergleichen Sie die Strukturen.

Um zu ermitteln, welche der fünf Vorhersagen dem Zielmodell am ähnlichsten ist, starten Sie

```
Tools -> MD/Ensemble Analysis -> Ensemble Match
```

und wählen Sie das Zielmodell als Referenz und die Vorhersagemodelle als Alternative aus. Geben Sie unter Parts to Match den Ausdruck für die Rückgratátome an (@n@ca@C) und klicken Sie auf OK. Im Ergebnisfenster sehen Sie die RMSD Werte zwischen den Vorhersagen und dem Zielmodell. Notieren Sie sich das Modell mit dem besten Wert und schalten Sie die Sichtbarkeit der anderen Vorhersagen aus. Notieren Sie sich die berechneten RMSD Werte oder behalten Sie das Ergebnisfenster bis zum ende der Übung offen.

Um die Ähnlichkeit per Residue zu quantifizieren, können einzelne RMSD Werte berechnet werden. Führen Sie dazu zunächst ein Sequenzalignment zwischen dem Zielmodell und der besten Vorhersage durch. Starten Sie im Alignmentfenster Structure -> Associations... und stellen Sie eine Assoziierung des Zielmodells und der besten Vorhersage jeweils mit sich selbst und keine Assoziierung (none) für alle anderen Modelle ein. Klicken Sie dann auf Headers -> RMSD: full.

Nun wird der per-Residue RMSD Wert als Histogramm über der Sequenz angezeigt. Über das Render/Select by Attribute Fenster können Sie nun die Strukturen nach berechnetem RMSD Wert (mavRMSDfull) färben und auswählen (wenn Sie den Reiter

Select verwenden). Lassen Sie sich die Seitenketten der stärker abweichenden Residues anzeigen und vergleichen Sie die Strukturen. Die ausgewählten Residues werden in den Alignmentfenstern Markiert. Schauen Sie sich das erste Alignment zwischen dem Ziel- und dem Templatemodell an und wählen Sie stark abweichende Residues in den Strukturen aus. Was beobachten Sie?

## Hinzunahme von Templates

Nach 1fca ist das nächstbeste durch BLAST gefundene Modell 1fdn. Laden Sie dieses und wiederholen Sie anschließend die Modelling Schritte (beginnend bei dem Sequenzalignment), wobei sie dieses Mal 1fdn zusätzlich als Template verwenden. Dies bedeutet dass Sie:

- Ein Sequenzalignment zwischen den drei Modellen 1dur, 1fca und 1fdn durchführen.
- Eine Modellierung von 1dur mit 1fca und 1fdn als Templates durchführen.

Vergleichen Sie die Qualitätsmetriken. Berechnen Sie auch hier die RMSD Werte zwischen Vorhersagen und Zielmodell mittels des Ensemble Match Tools und vergleichen Sie die Werte mit den vorherigen Ergebnissen.