

Übungen zur Vorlesung Algorithmische Bioinformatik

Freie Universität Berlin, WS 2014/15

Martin Vingron · Juliane Perner · Annkatrin Bressin

Blatt 14 · Ausgabe am 26.01.2015

Abgabe am 02.02.2015 vor Beginn der Vorlesung

Name:

Matrikelnummer:

Übungsgruppe:

Aufgabe 1 (10 Punkte; Theorie). Nehmen Sie an, sie haben folgende Torsionswinkel (ϕ, ψ) einer Sekundärstruktur gegeben:

$(-70, -20)$, $(-72, 60)$, $(-70, 120)$, $(-60, 170)$, $(-65, 125)$,
 $(-100, 45)$, $(-100, -65)$, $(-105, -66)$, $(-100, -60)$

1. In der Vorlesung haben Sie eine graphische Veranschaulichung kennengelernt, mit der Sie anhand der Torsionswinkel abschätzen können, ob es sich um eine realistische Struktur handelt. Wie heißt sie? Beschreiben Sie kurz das Prinzip.
2. Welche der oben aufgelisteten Torsionswinkelpaare könnten in einer α -Helix oder in einem β -Strang liegen?

Aufgabe 2 (20 Punkte; Theorie). Gegeben seien zwei 2D-Strukturen A und B mit den folgenden Koordinaten:

A: $(1, 4)$, $(4, 1)$, $(4, 4)$
B: $(0, 0)$, $(2, 0)$, $(3, 2)$

Die Spalten seien dabei die äquivalenten Punkte in den Strukturen.

1. Berechnen Sie die RMSD zwischen den beiden Strukturen.
2. Berechnen Sie die Massenzentren der beiden Strukturen und verschieben Sie sie so, dass die Massenzentren auf dem Koordinatenursprung liegen. Stellen Sie die neuen Koordinaten in einem Plot graphisch dar.
3. Berechnen Sie nun erneut die RMSD.
4. Welche Möglichkeiten gibt es um die RMSD weiter zu verringern?

Aufgabe 3 (30 Punkte; Programmieren). Wir möchten nun die Strukturen zweier Proteine¹² vergleichen.

1. Schauen Sie im Kopf der PDB-Dateien oder mit Hilfe der Identifier nach um welches Protein es sich handelt. Was ist der Unterschied zwischen den beiden PDBs?
2. Implementieren Sie eine Funktion, die den Namen einer PDB-Datei erhält und aus dieser Datei die Koordinaten der C_α -Atome einliest.
3. Sie möchten nun das Massenzentrum des Moleküls zum Koordinatenursprung verschieben. Implementieren Sie dazu eine Funktion die diese Translation auf den eingelesenen Daten ausführt.
4. Vergleichen Sie nun die beiden Strukturen in dem sie die RMSD zwischen den C_α -Atome der gleichen Aminosäuren berechnen.

Aufgabe 4 (40 Punkte; Theorie). In der Vorlesung wurde ein Algorithmus zum Finden des besten Überlapps zweier Strukturen basierend auf Double-Dynamic Programming vorgestellt. Gegeben seien zwei Peptidsequenzen ABC und DEF mit folgenden Koordinaten

	x	y	z			x	y	z
A	1	0	0		D	0	1	0
B	0	1	0		E	0	1	1
C	0	1	1		F	1	0	0

1. Berechnen Sie alle neun Low-level-Matrizen. Verschieben Sie dazu das zweite Molekül, sodass die fixierten Residuen auf der gleichen Position liegen. Nehmen Sie außerdem folgende Scoring-Funktion an:

$$S = \left(\frac{1}{d} - \frac{1}{1.75} \right) * 10$$

Dabei ist d die Distanz zwischen den Atomen. Falls $d = 0$, setzen Sie den maximalen Score S auf 5. Gaps werden nicht bestraft.

2. Berechnen Sie die High-level-Matrix und führen Sie den dynamischen Programmieralgorithmus aus. Welches ist der beste Überlapp?

¹Material1: <http://www.molgen.mpg.de/Algorithmische-Bioinformatik-WS1415/u14/1A3N>

²Material2: <http://www.molgen.mpg.de/Algorithmische-Bioinformatik-WS1415/u14/1GZX>