ITA

Scaricate il file pdb 1yzb dal Protein Data Bank.

1yzb è un file con una traiettoria di 10 steps di una proteina. Potete visualizzarla con vmd

vmd 1yzb.pdb

Ciascuno step separato dagli altri dalla riga ENDMDL.

Per ciascuno degli esercizi ripotati sotto, dovete produrre degli script che riescano a rispondere ai quesiti. Siate ordinati, e per ciascun esercizio create una cartella con lo script che svolge l'operazione e il file sorgente (1yzb.pdb).

1) Prendere il pdb e ottenere un file contenente solo le righe con le coordinate degli atomi, eliminando tutti i commenti. ATTENZIONE A CONSERVARE TUTTI I MODELLI SEPARATI DA ENDMDL. Da ora in poi per tutti gli esercizi successivi considerate come input quel file che otterrete. (Sarà  utile il comando grep)

2) Scrivere uno script che riconosca il numero di campi di testo contenuti nel pdb, estragga il numero e il nome degli atomi e li salvi in un file di testo (grep e awk)

3) Trasformare la traiettoria pdb in una traiettoria xyz dove il file xyz è fatto come segue:

"

NUMERO ATOMI TOTALE

TITOLO

ATOMO X Y Z

NUMERO ATOMI TOTALE

TITOLO

ATOMO X Y Z

"

Esempio di una traiettoria di 2 step con 2 atomi

"

2

prova

N 13.567 12.589 13.245

O 14.234 15.325 11.789

2

prova

N 13.567 12.589 13.245

O 14.234 15.325 11.789

"

Controllate che tutto sia ok visualizzando la traiettoria con vmd

4) Estrarre la traiettoria del backbone della proteina, e salvarla in un file xyz con tutti gli step

5) Estrarre dal file le POSIZIONI MEDIE del backbone e scriverle in un file xyz. Confrontate il file con la traiettoria su vmd.

6) Calcolare l'rmsd di ciascuno step della traiettoria rispetto alla configurazione media che avete calcolato, considerando SOLO I CARBONI ALPHA (chiamati CA).

ENG

Download the 1yzb pdb file from the Protein Data Bank.

1yzb is a file with a 10 steps protein trajectory. You can view it with vmd

vmd 1yzb.pdb

Each step is separated from the others by the ENDMDL line.

For each exercise, you have to produce bash scripts that can answer the questions.

1) Take the PDB and get a file containing only the lines with the coordinates of the atoms, deleting all comments. ATTENTION TO KEEP ALL MODELS SEPARATED BY ENDMDL. From now on this output file will be considered as input for all other exercises. (Grep command will be useful)

2) Write a script that recognizes the number of text fields contained in the pdb, extract the number and name of the atoms and save them in a text file (grep and awk)

3) Change format from pdb to xyz

xyz format is as follows:

"

TOTAL ATOMIC NUMBER

TITLE

ATOMO X Y Z

TOTAL ATOMIC NUMBER

TITLE

ATOMO X Y Z

"

Example of a 2-step trajectory with 2 atoms

"

2

test

N 13,567 12,589 13,245

O 14,234 15,325 11,789

2

test

N 13,567 12,589 13,245

O 14,234 15,325 11,789

"

Check that everything is ok by viewing the trajectory with vmd

4) Extract the trajectory of the protein backbone, and save it in an xyz file with all steps

5) Extract from the file the AVERAGE POSITIONS of the backbone and write them in an xyz file. Compare the file with the trajectory on vmd.

6) Calculate the total rmsd throughout the trajectory with respect to the average configuration that you have calculated in exercise 5. Considering ONLY THE ALPHA CARBONS (called CA).