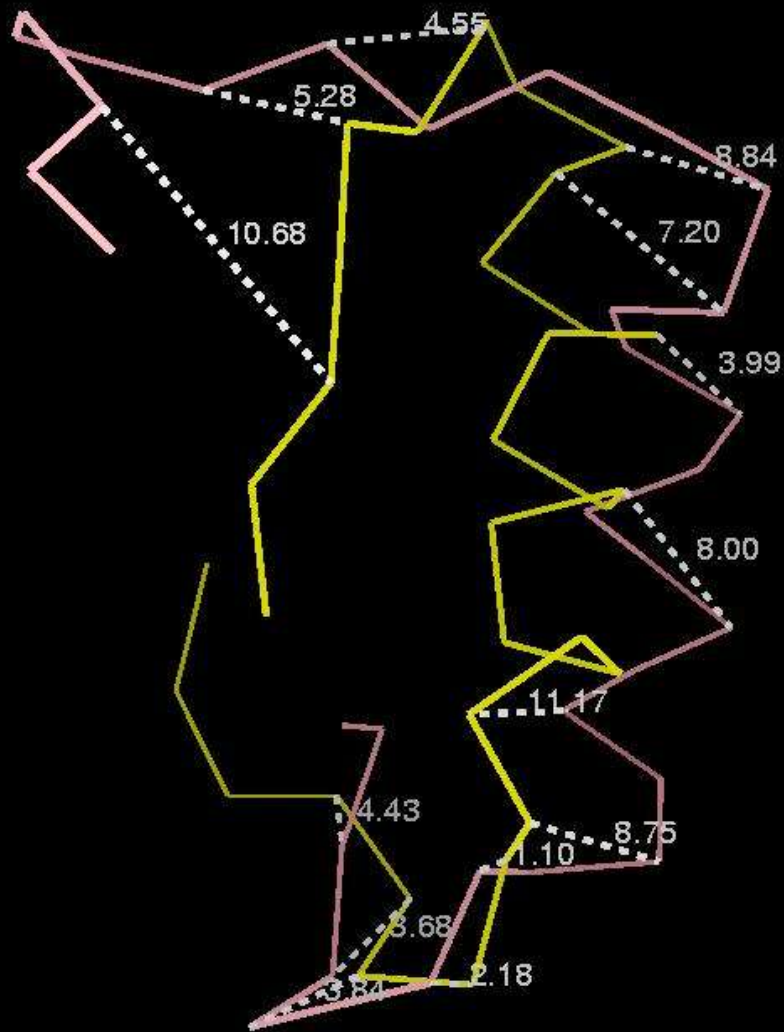


1. Mesurer la similitude entre 2 formes

2. Superposition “optimale” de deux structures



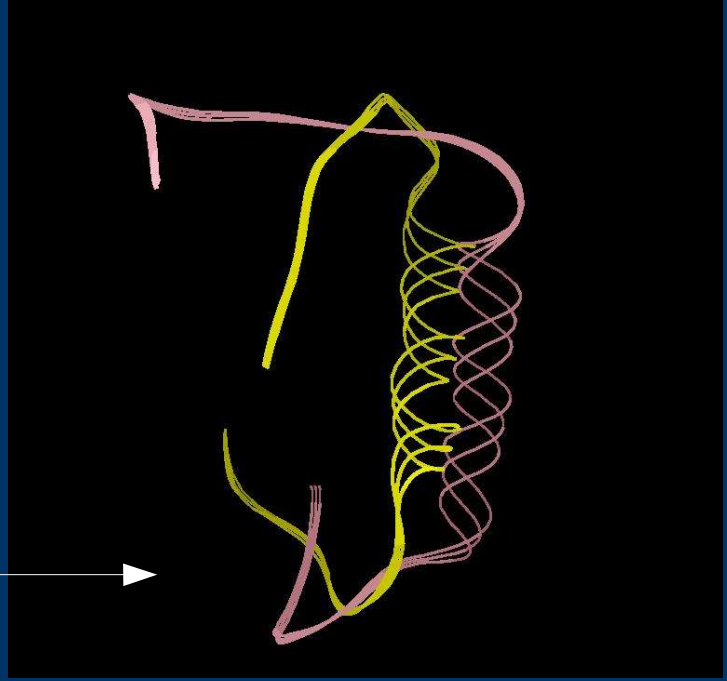
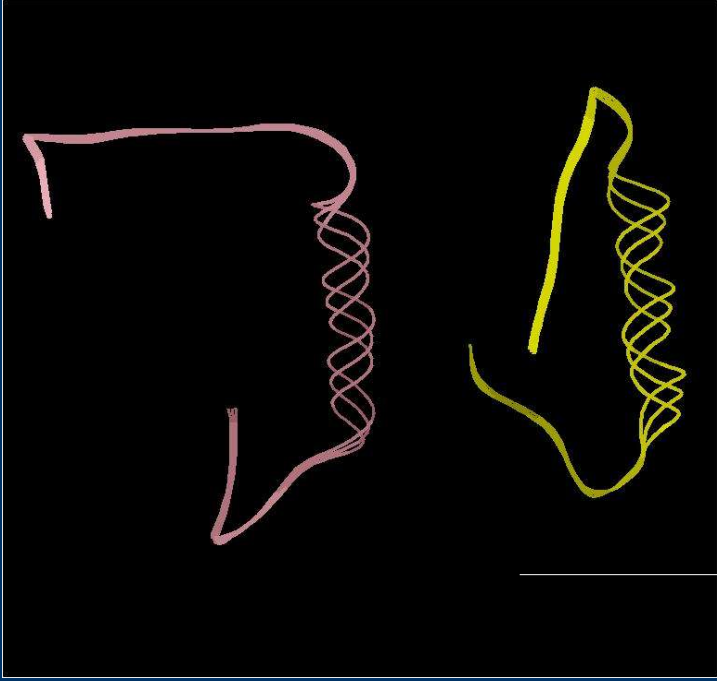


- **cRMSd**

$$\text{cRMSd} = \sqrt{\sum_i d_i^2 / n}$$

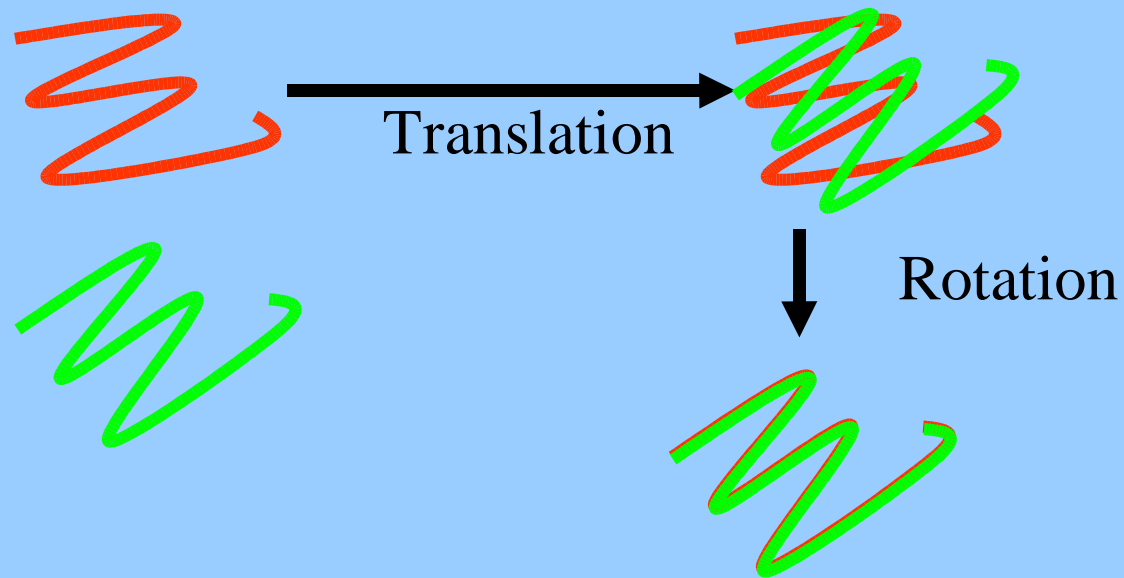
n = nombre d'atomes

**d_i = distance entre 2 atomes i correspondant
dans les 2 structures**



- **Superposition de deux structures**

Translation and Rotation



- **Superposition de deux structures**

Algorithmes (cRMSd):

- **Changement de repère**

(Vecteurs propres de la matrice d'inertie)

- **Moindres carrés**

(Sipl M., Stegbuchner H., *Computers and Chemistry* 15:73-78, 1991)

<ftp://@ftp.came.sbg.ac.at/pub/SuperPos>

- **Quaternions**

(Zuker & Somorjai, *Bulletin of Mathematical Biology*, 51: 55-78, 1989.)



Limites du cRMSd

→ Tous les atomes sont traités de manière équivalente

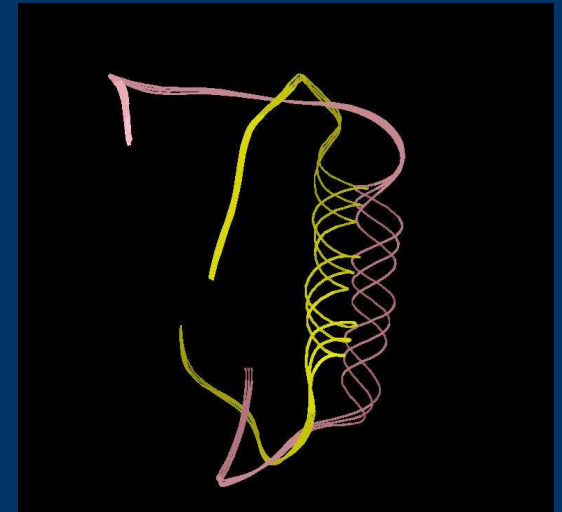
Alternatives (W. Taylor(1999), *Protein Science*, 8: 654-665.):

RMSd pondéré

RMSd limité aux paires se superposant le mieux

→ La signification du RMSd dépend du nombre de comparaisons

Alternative: (Betancourt M., Skolnick (2001) J., *Biopolymers*, 59:305-309)



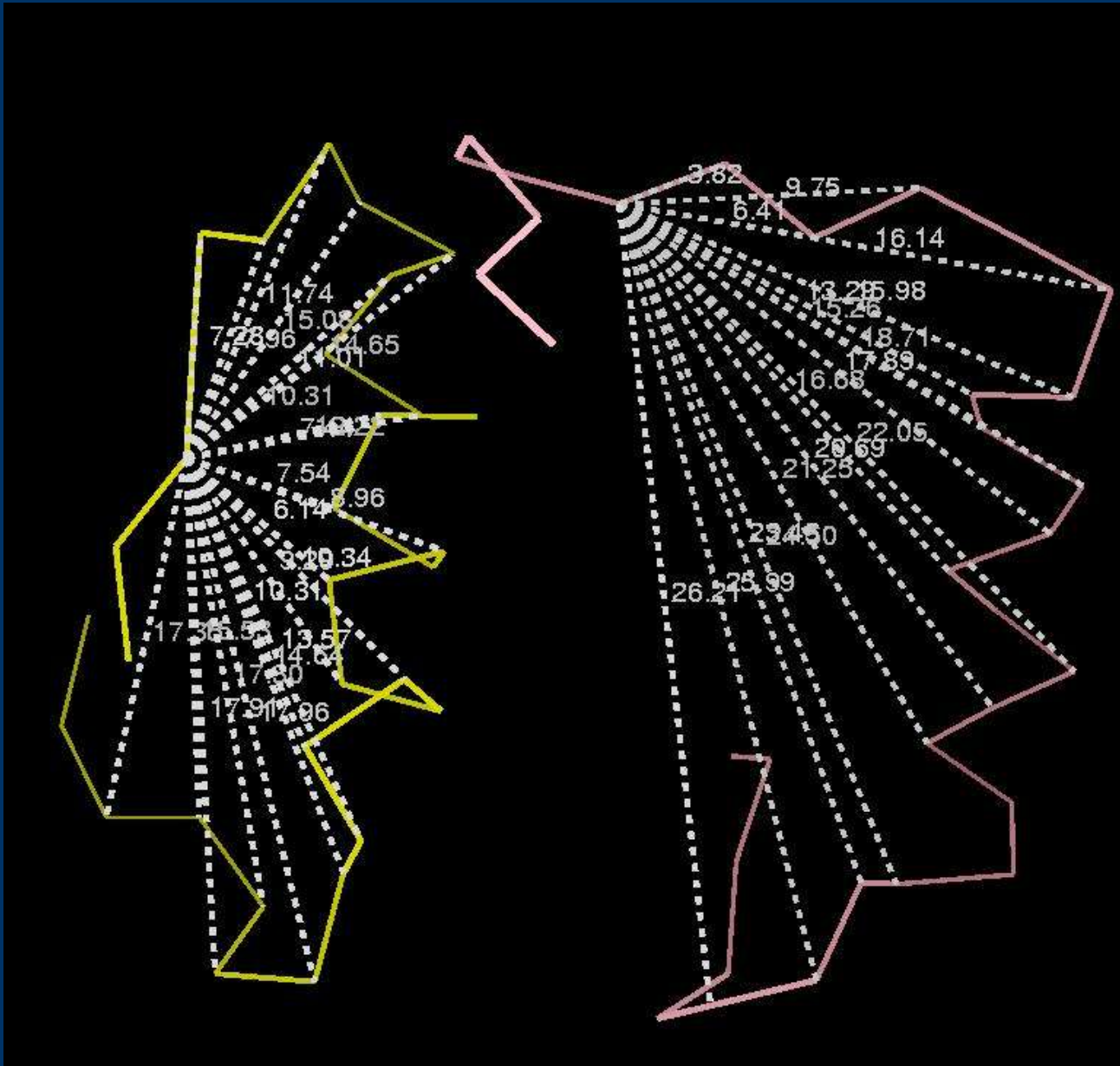
- **dRMSd**

$$dRMSd = \sqrt{2/n(n-1) \sum_{i=2}^n \sum_{j=1}^{i-1} (d_{(ij)}^P - d_{(ij)}^Q)^2}$$

n = nombre d'atomes

d_{ij} = distance entre 2 atomes i et j d'une même structure

P, Q: les structures à comparer





- **Outils de superposition de deux structures sachant un alignement**

→ **ProFit**

(<http://www.bioinf.org.uk/>)

Nécessite:

- un fichier d'alignement (format **NBRF**)
 - un fichier des coordonnées correspondant (Format **PDB**)
-
-

Recherche de l'alignement structural “optimal”

- **Le problème de base :**
 - **On a 2 structures protéiques de longueur non nécessairement identiques**
 - **On cherche à les superposer au mieux**

==> On recherche l'alignement structural

Alignement structural

- **Problème complexe**
- **Nombreuses approches**
 - ★ **Topologie (structure secondaire)**
 - **TOP** (<http://bioinfo1.mbfys.lu.se/TOP/webtop.html>)
 - **K2** (<http://zlab.bu.edu/k2/index.shtml>)
 - ★ **Similitudes locales**
 - **CE** (http://cl.sdsc.edu/ce/ce_align.html)
 - **LGA** (<http://predictioncenter.llnl.gov/local/lga/lga.html>)
 - **ProFit**

Alignement structural

- **Limites:**

- **Signification de l'alignement pour des structures dissemblables ?**

==> significativité du critère de similitude

- **On peut avoir des régions semblables et d'autres dissemblables**

==> alignement local versus alignement global



Recherche de structures similaires

- **Le problème complet :**
 - **On a N structures protéiques**
 - **On cherche à identifier les régions similaires**

**==> On recherche des alignements
structuraux locaux qui satisfont un critère
de similitude**

Approches les plus connues (1)

- **DALI – FSSP** (<http://www.ebi.ac.uk/dali/Interactive.html>):
 - Basé sur un critère dRMSd
 - Recherches de similitudes structurales parmi l'ensemble des structures connues par monte-Carlo
 - Chaque protéine est décrite dans un banque

Approches les plus connues (2)

- **CE** (<http://cl.sdsc.edu/ce.html>):
 - **Basé sur un critère complexe**
 - **Recherches de similitudes structurales**
parmi l'ensemble des structures connues avec
un algorithme complexe
 - **Chaque protéine est décrite dans un banque**

Une autre approche (parmi N)

- **SA-PSimSearch**

(*<http://condor.urbb.jussieu.fr/cgi-bin/SA-PSimSearch-cgi.py>*):

- **basé sur un encodage des protéines dans un alphabet structural**
 - **Recherches de similitudes structurales basée sur les techniques classiques d'alignement de séquences**
-
-

