

# Prédiction de propriétés structurales

- **A partir de la séquence**
  - Prédiction de l'accessibilité au solvant
  - Localisation des ponts disulfure
  - Localisation des domaines de repliement
  - Localisation des sites fonctionnels
- **A partir de la structure**
  - Identification des zones d'interaction
  - Identification des sites fonctionnels
  - Mutabilité

# Prédiction de l'accessibilité au solvant

→ basée sur un apprentissage :

- des propensités des acides aminés à être enfouis.
- de descripteurs globaux des protéines

→ utilise une fenêtre sur la séquence

(i.e. la prédiction en un site dépend des acides aminés alentour)

NETASA (réseau de neurones) : <http://www.netasa.org/>

PredAcc (perceptron): <http://condor.urbb.jussieu.fr/logiciels/PredAccCfg.html>

---

---

# Localisation des ponts disulfures

## ★ Prédiction de l'état d'oxydation des cystéines

→ basée sur un apprentissage :

- de descripteurs globaux des protéines
- de la séquence flanquant les cystéines
- de l'état d'oxydation dans des séquences homologues

CysREDOX: *seems abandoned*

CysPred (Neural net): <http://gpcr.biocomp.unibo.it/predictors/cyspred/>

CysState (perceptron): <http://condor.urbb.jussieu.fr/cgi-bin/CysState-cgi.py>

SVM Cystéines : <http://cassandra.dsi.unifi.it/cysteines/>

## ★ Prédiction de l'appariement des cystéines

SVM Cystéines : <http://cassandra.dsi.unifi.it/cysteines/>

DCON : [http://gpcr.biocomp.unibo.it/cgi/predictors/cys-cys/pred\\_dconcgi.cgi](http://gpcr.biocomp.unibo.it/cgi/predictors/cys-cys/pred_dconcgi.cgi)

# Localisation des domaines de repliement

- ★ **Homologie avec les séquences de domaines identifiés**  
(**PFAM:** *<http://www.sanger.ac.uk/Software/Pfam/>*)  
(**ProDOM:** *<http://protein.toulouse.inra.fr/prodom/current/html/home.php>*)

DomPred: *<http://bioinf.cs.ucl.ac.uk/dompred/>*



# Localisation des sites fonctionnels

## ★ Homologie avec des motifs fonctionnels identifiés

ProSite: *<http://us.expasy.org/prosite/>*

ELM: *<http://elm.eu.org/>*

interPro:*<http://www.ebi.ac.uk/interpro/>*

# A partir de la structure:

- **Identification des domaines :**

**Protein domain server :**

*[http://www.bmm.icnet.uk/~domains/assign\\_by\\_form.html](http://www.bmm.icnet.uk/~domains/assign_by_form.html)*

**domId** : *<http://bioinfo1.mbfys.lu.se/Domid/domid.html>*

- **Prédiction des zones d'interaction**

**interPrets :**

*<http://www.russell.embl-heidelberg.de/people/patrick/interprets/interprets.html>*

**isPred** : *[http://gpcr.biocomp.unibo.it/cgi/predictors/pp/pred\\_pp.cgi](http://gpcr.biocomp.unibo.it/cgi/predictors/pp/pred_pp.cgi)*



# A partir de la structure:

- **Identification des sites fonctionnels**

Methodes couplant structure et homologie  
(ex: Yao et al. J.Mol.Biol.(2003)326:255-261)

THEMATICS (Ondrechen et al., PNAS (2001) 98:12473-12478)

- **Impact de mutations ponctuelles**

FoldX : <http://fold-x.embl-heidelberg.de:1100/cgi-bin/main.cgi>

popMusic: <http://babylone.ulb.ac.be/popmusic/>

