Bioinformatics III

Analysis and prediction of 3D macromolecule structures

lecture 2
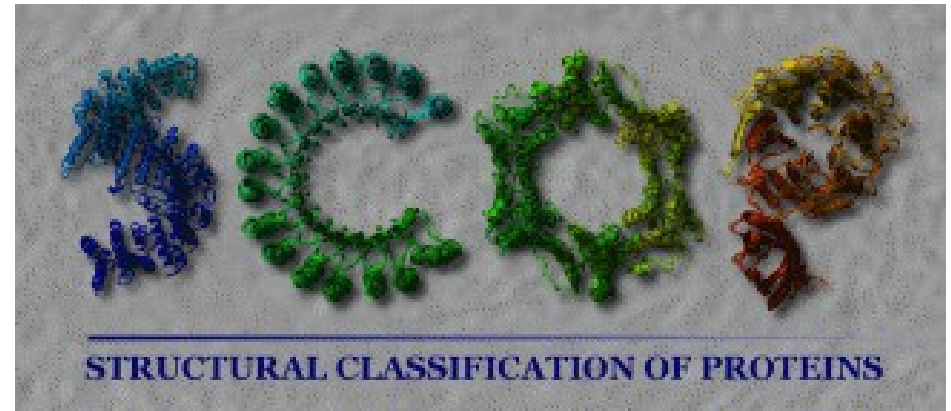classification of protein structures

Saulius Gražulis
2022-02-15

Classifications of protein structures

- SCOP classification
- CATH classification
- FSSP/Dali classification
- ANN Autoencoders?

SCOP{2,e} classification

- Human curated
  (Alexey G. Murzin, et al.,
  Cyrus Chothia)

- Hierarchy:
  - Class
  - Fold
  - Superfamily
  - Family

http://scop.mrc-lmb.cam.ac.uk/scop/
http://scop2.mrc-lmb.cam.ac.uk/
https://scop.berkeley.edu/

SCOPe classes

Classes in SCOPe 2.08:

1. a: All alpha proteins [46456] (290 folds)
2. b: All beta proteins [48724] (180 folds)
3. c: Alpha and beta proteins (a/b) [51349] (148 folds)|
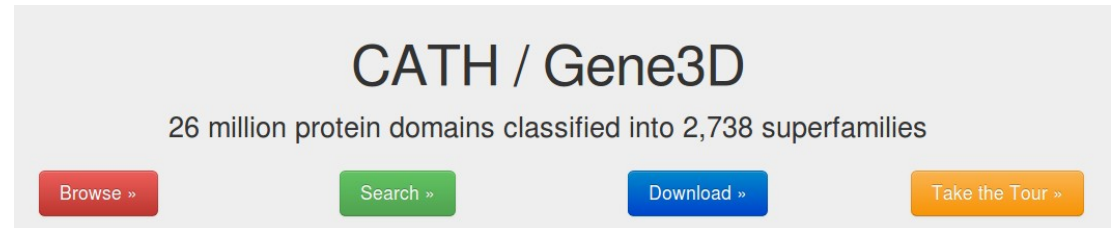4. d: Alpha and beta proteins (a+b) [53931] (396 folds)
5. e: Multi-domain proteins (alpha and beta) [56572] (74 folds)
6. f: Membrane and cell surface proteins and peptides [56835] (69 folds)
7. g: Small proteins [56992] (100 folds)
8. h: Coiled coil proteins [57942] (7 folds)
9. i: Low resolution protein structures [58117] (25 folds)
10. j: Peptides [58231] (151 folds)
11. k: Designed proteins [58788] (44 folds)
12. l: Artifacts [310555] (1 fold)

https://scop.berkeley.edu/

CATH classification

- (Almost) fully automatic

- Hierarchy:

  - Class

  - Architecture

  - Topology

  - Homologuos superfamily

CATH / Gene3D
26 million protein domains classified into 2,738 superfamilies

Browse »   Search »   Download »   Take the Tour »

http://www.cathdb.info/

(Sillitoe 2021 https://doi.org/10.1093/nar/gkaa1079)

Dali/FSSP

https://www.ebi.ac.uk/msd-srv/ssm/

- Original link dead?

(Holm 1997 https://doi.org/10.1093/nar/25.1.231)