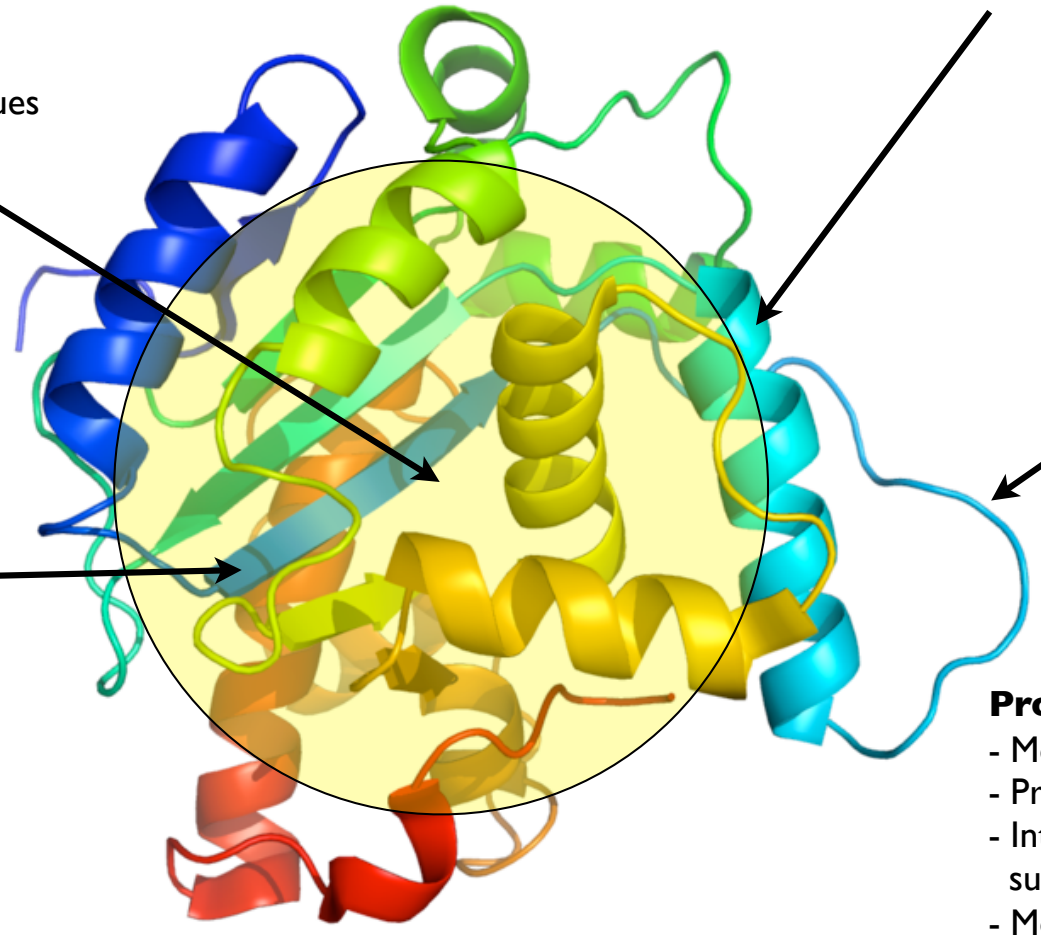
CHEM 436 / 630

Molecular modelling of proteins

Winter 2018 Term

Instructor:
Guillaume Lamoureux
Concordia University, Montréal, Canada

Protein structure



Alpha helix

Protein core
- Mostly hydrophobic residues
- Densely packed

Loop

Beta strand (forming a beta sheet)

Protein surface
- Mostly hydrophilic residues
- Protein-protein interfaces
- Interactions with membrane, substrate, etc.
- More dynamic than the core (side chain motion, loop motion, hinge motion, etc.)

Source: **Wikipedia**
http://upload.wikimedia.org/wikipedia/commons/e/e6/Spombe_Pop2p_protein_structure_rainbow.png

Finding the structure of a protein from its sequence

Levinthal's Paradox
(Why do proteins fold at all?)

Let's say that each amino acid of a 100-residue long protein can be in one of three conformations: extended, helical, or loop.

100 AAs
3 conformations each

If each AA adopts one of those conformations independently of it neighbors, the whole protein has 3^{100} possible conformations.

$3^{100} \approx 5\times10^{47}$ conformations

Let's presume that each conformation converts into any other one at the rate of 10^{14} s^{-1} (one isomerization every 0.01 ps).

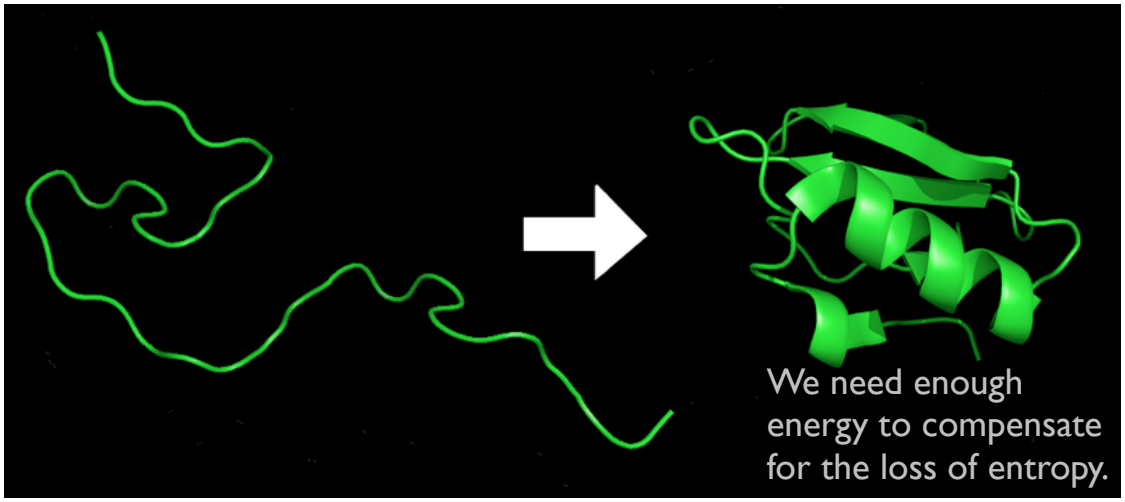
$5\times10^{47} / (10^{14}\ s^{-1}) = 5\times10^{33}$ s
$= 1.6\times10^{33}$ y
to visit all conformations

Age of the universe : **1.4×10^{10} y**

Proteins usually fold in a few milliseconds or microseconds.

Small segments of the sequence rapidly "nucleate" a specific secondary structure. Once these segments are "done", there is not much "random exploration" left to do.

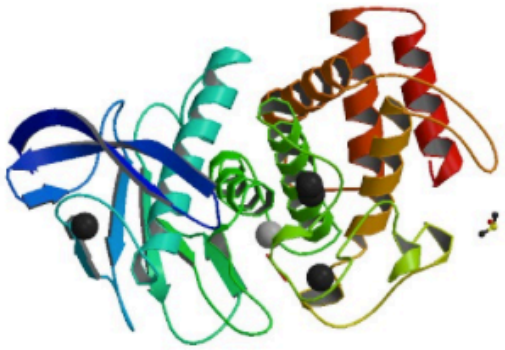
Why are protein folding?
What make proteins "collapse" in a given 3D structure?



- Hydrophobic effect
- Hydrogen bonds
- Salt bridges
- Metal ions (Ca^{2+}, Mg^{2+}, Zn^{2+}, etc.)
- etc.

Source: **Wikipedia**
http://en.wikipedia.org/wiki/Protein_folding

Biological Assembly 1



3D View: Structure | Electron Density | Ligand Interaction

Standalone Viewers
Protein Workshop | Ligand Explorer

Standalone Viewers
Protein Workshop | Ligand Explorer

Global Symmetry: Asymmetric - C1
Global Stoichiometry: Monomer - A

Biological assembly 1 assigned by authors and generated by PISA (software)

Display Files | Download Files

1LNF

A structural analysis of metal substitutions in thermolysin

DOI: 10.2210/pdb1LNF/pdb

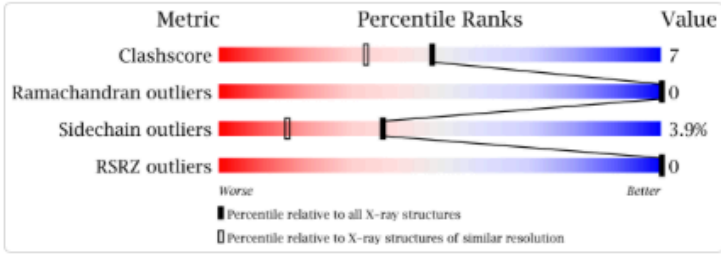
Classification: HYDROLASE (METALLOPROTEASE)
Organism(s): Bacillus thermoproteolyticus

Deposited: 1994-05-13 **Released:** 1995-05-08
Deposition Author(s): Holland, D.R., Matthews, B.W.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 1.7 Å

wwPDB Validation 3D Report | Full Report



This is version 1.5 of the entry. See complete history.

Literature

Download Primary Citation

Structural analysis of zinc substitutions in the active site of thermolysin.

Holland, D.R., Hausrath, A.C., Juers, D., Matthews, B.W.
(1995) Protein Sci. **4**: 1955-1965

PubMed: 8535232 Search on PubMed | Search on PubMed Central
DOI: 10.1002/pro.5560041001
Primary Citation of Related Structures: 1LNA, 1LNB, 1LNC, 1LND, 1LNE

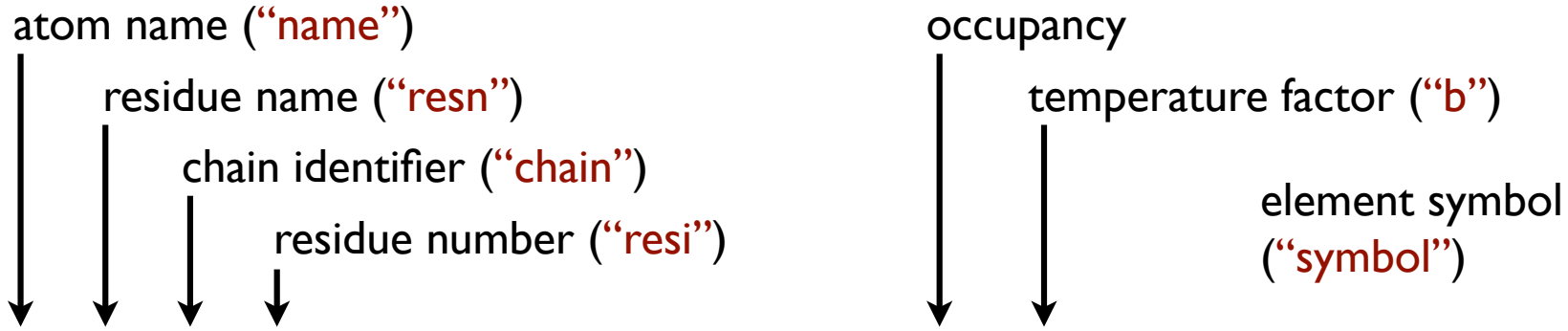
PDB file format

(details at http://www.wwpdb.org/documentation/file-format)

```
HEADER    HYDROLASE (METALLOPROTEASE)                13-MAY-94   1LNF
TITLE     A STRUCTURAL ANALYSIS OF METAL SUBSTITUTIONS IN THERMOLYSIN
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: THERMOLYSIN;
COMPND   3 CHAIN: E;
COMPND   4 EC: 3.4.24.27;
COMPND   5 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: BACILLUS THERMOPROTEOLYTICUS;
SOURCE   3 ORGANISM_TAXID: 1427
KEYWDS    HYDROLASE (METALLOPROTEASE)
EXPDTA    X-RAY DIFFRACTION
AUTHOR    D.R.HOLLAND,B.W.MATTHEWS
REVDAT   4   25-AUG-09 1LNF    1       SOURCE
REVDAT   3   14-JUL-09 1LNF    1       REMARK
REVDAT   2   24-FEB-09 1LNF    1       VERSN
REVDAT   1   08-MAY-95 1LNF    0
JRNL        AUTH   D.R.HOLLAND,A.C.HAUSRATH,D.JUERS,B.W.MATTHEWS
JRNL        TITL   STRUCTURAL ANALYSIS OF ZINC SUBSTITUTIONS IN THE
JRNL        TITL 2 ACTIVE SITE OF THERMOLYSIN.
JRNL        REF    PROTEIN SCI.                  V.   4  1955 1995
JRNL        REFN                   ISSN 0961-8368
JRNL        PMID   8535232
REMARK   1
REMARK   1 REFERENCE 1
REMARK   1  AUTH   D.R.HOLLAND,D.E.TRONRUD,H.W.PLEY,K.M.FLAHERTY,
REMARK   1  AUTH 2 W.STARK,J.N.JANSONIUS,D.B.MCKAY,B.W.MATTHEWS
REMARK   1  TITL   STRUCTURAL COMPARISON SUGGESTS THAT THERMOLYSIN AND
```

PDB file format

(details at http://www.wwpdb.org/documentation/file-format)




```
ATOM      1  N   ILE E   1       9.198  51.923  -1.205  1.00 56.09           N
ATOM      2  CA  ILE E   1       8.244  52.190  -0.120  1.00 34.74           C
ATOM      3  C   ILE E   1       7.870  53.687   0.082  1.00 41.10           C
ATOM      4  O   ILE E   1       8.699  54.583  -0.005  1.00 37.75           O
ATOM      5  CB  ILE E   1       8.733  51.578   1.191  1.00 39.04           C
ATOM      6  CG1 ILE E   1       8.149  50.191   1.418  1.00 26.23           C
ATOM      7  CG2 ILE E   1       8.402  52.483   2.377  1.00 61.96           C
ATOM      8  CD1 ILE E   1       8.191  49.808   2.899  1.00100.00           C
ATOM      9  N   THR E   2       6.618  53.984   0.430  1.00 35.26           N
ATOM     10  CA  THR E   2       6.254  55.377   0.695  1.00 34.94           C
ATOM     11  C   THR E   2       5.978  55.528   2.171  1.00 35.54           C
ATOM     12  O   THR E   2       5.159  54.798   2.727  1.00 43.29           O
ATOM     13  CB  THR E   2       5.025  55.832  -0.111  1.00 44.88           C
ATOM     14  OG1 THR E   2       5.277  55.636  -1.477  1.00 44.03           O
ATOM     15  CG2 THR E   2       4.750  57.309   0.150  1.00 48.14           C
ATOM     16  N   GLY E   3       6.669  56.455   2.825  1.00 19.43           N
ATOM     17  CA  GLY E   3       6.467  56.590   4.234  1.00 12.80           C
ATOM     18  C   GLY E   3       7.170  57.813   4.752  1.00 22.88           C
ATOM     19  O   GLY E   3       7.464  58.772   4.026  1.00 29.00           O
ATOM     20  N   THR E   4       7.338  57.812   6.033  1.00 14.24           N
ATOM     21  CA  THR E   4       7.906  58.949   6.694  1.00 11.42           C
```

Resolution (Å)	Meaning
> 4.0	Individual coordinates meaningless. Secondary structure elements can be determined.
3.0 – 4.0	Fold possibly correct, but errors are very likely. Many sidechains placed with wrong rotamer.
2.5 – 3.0	Fold likely correct except that some surface loops might be mismodelled. Several long, thin sidechains (Lys, Glu, Gln, etc.) and small sidechains (Ser, Val, Thr, etc.) likely to have wrong rotamers.
2.0 – 2.5	As 2.5–3.0, but number of sidechains in wrong rotamer is considerably less. Many small errors can normally be detected. Fold normally correct and number of errors in surface loops is small. Water molecules and small ligands become visible.
1.5 – 2.0	Few residues have wrong rotamer. Many small errors can normally be detected. Folds are rarely incorrect, even in surface loops.
0.5 – 1.5	In general, structures have almost no errors at this resolution. Individual atoms in a structure can be resolved. Rotamer libraries and geometry studies are made from these structures.

Source: **Wikipedia**
http://en.wikipedia.org/wiki/Resolution_(electron_density)

Protein structure

Hydrogen bonds
Don't trust the positions of the hydrogen atoms in the PDB. Except for very high resolution structures (< 1.0 Å), they are put in automatically —and often incorrectly.

Structural metals (Ca^{2+}, Mg^{2+}, Zn^{2+}, etc.)
Stabilize loops, enhance protein-protein (or domain-domain) interactions.

Catalytic metals
At or near the active site. Usually coordinated by Asp, Glu, His, or Cys.

Disulfide bonds
Between two Cys. Enhance the stability of proteins in oxidative environment.

Cofactors
Ligands
Water molecules
etc.

Buried water molecules:
https://dx.doi.org/10.1002/pro.5560030808

Water molecules directly participating in the function of a K^+ channel:
https://dx.doi.org/10.1038/nature12395