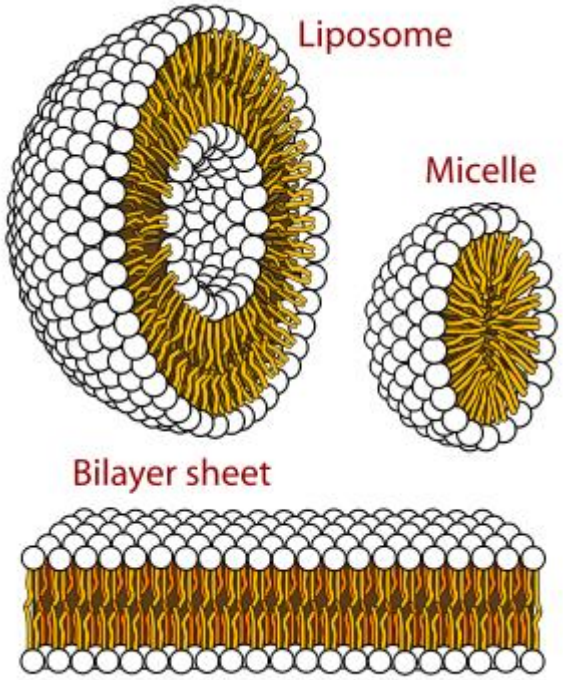
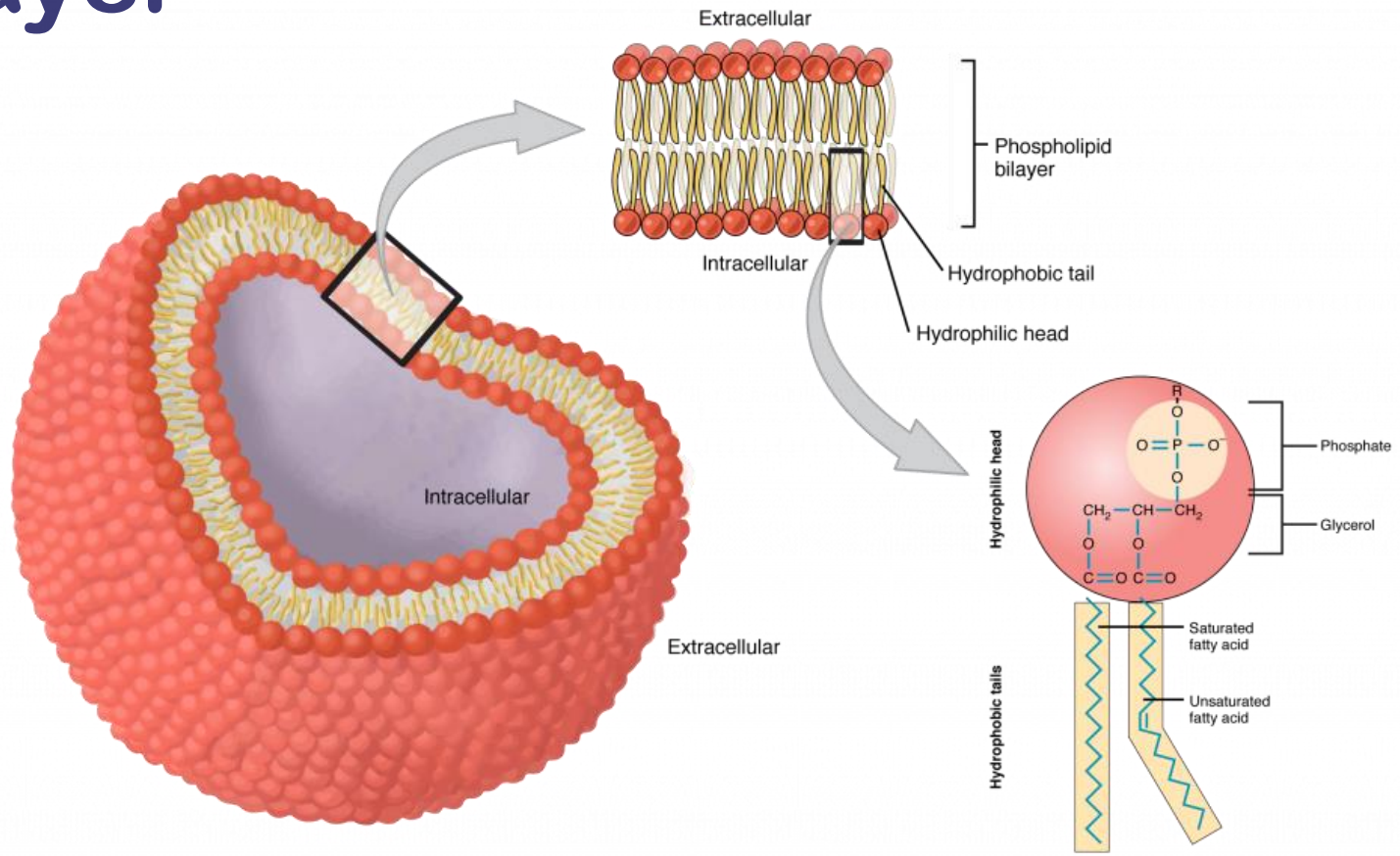
# Macromolecules
# Transmembrane proteins

Csaba Magyar, Institute of Enzymology,
Research Centre for Natural Sciences
2019 December

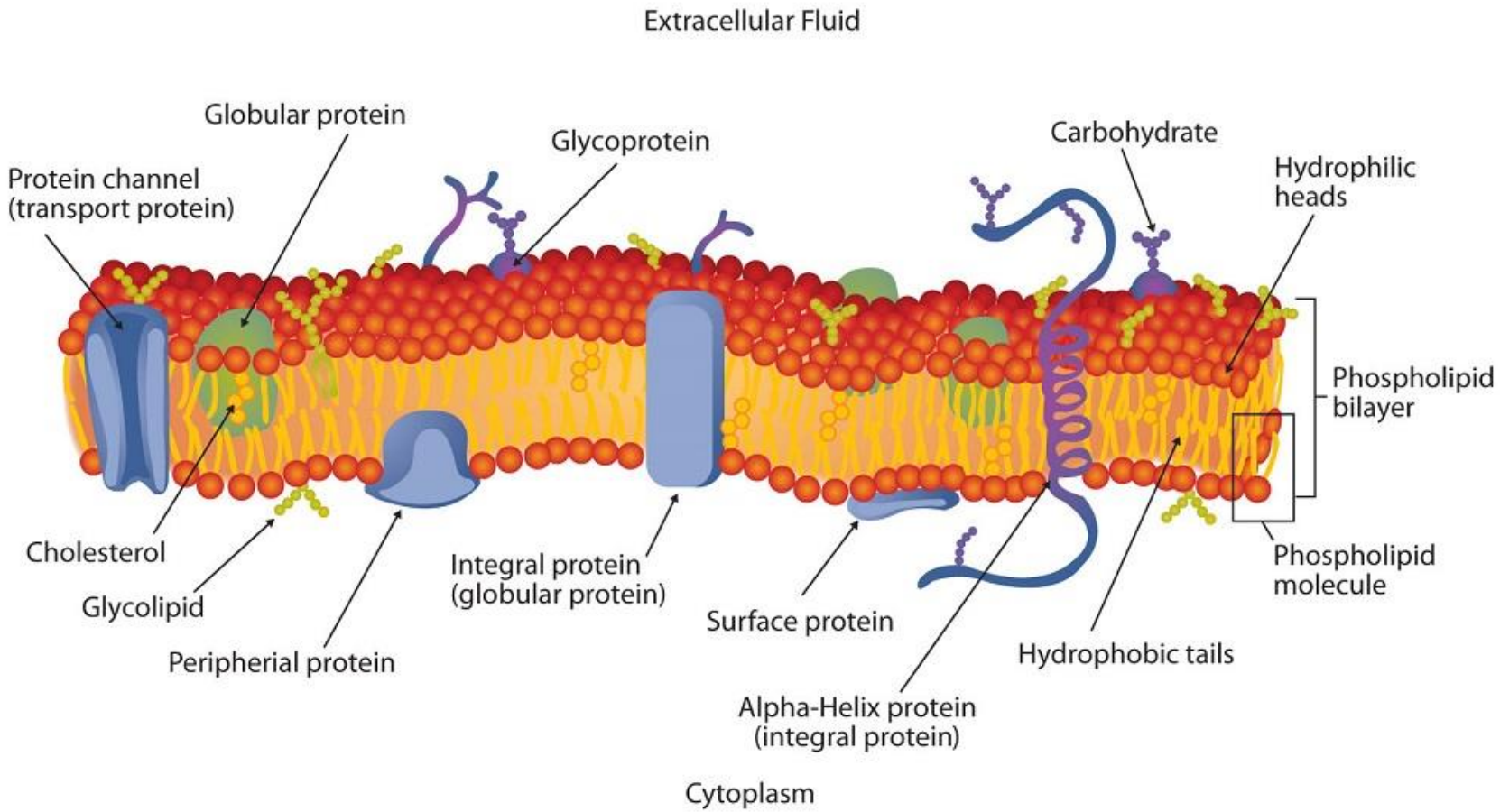
# Phospholipid bilayer




**Separation**
**Polarity**

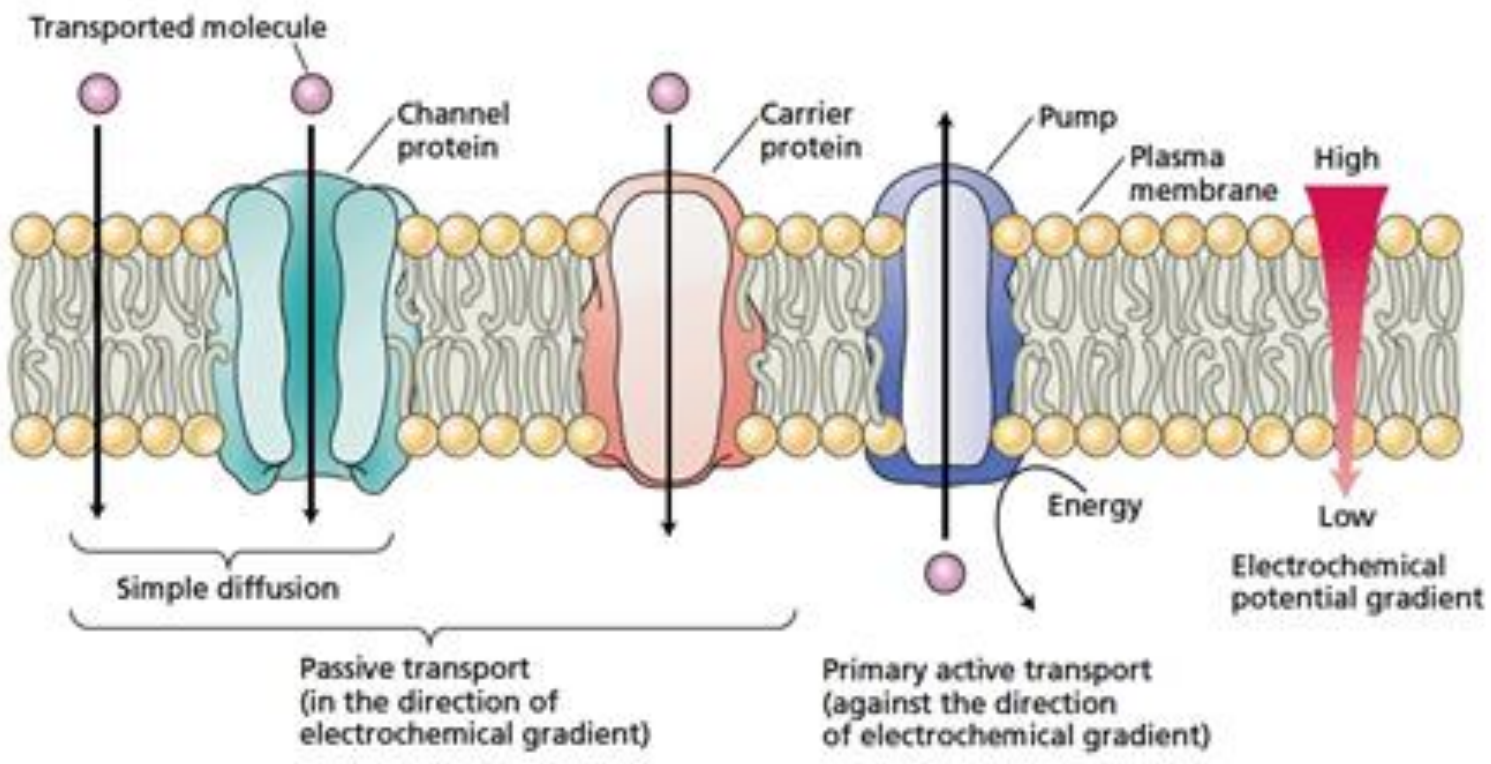
# Function of lipid bilayers



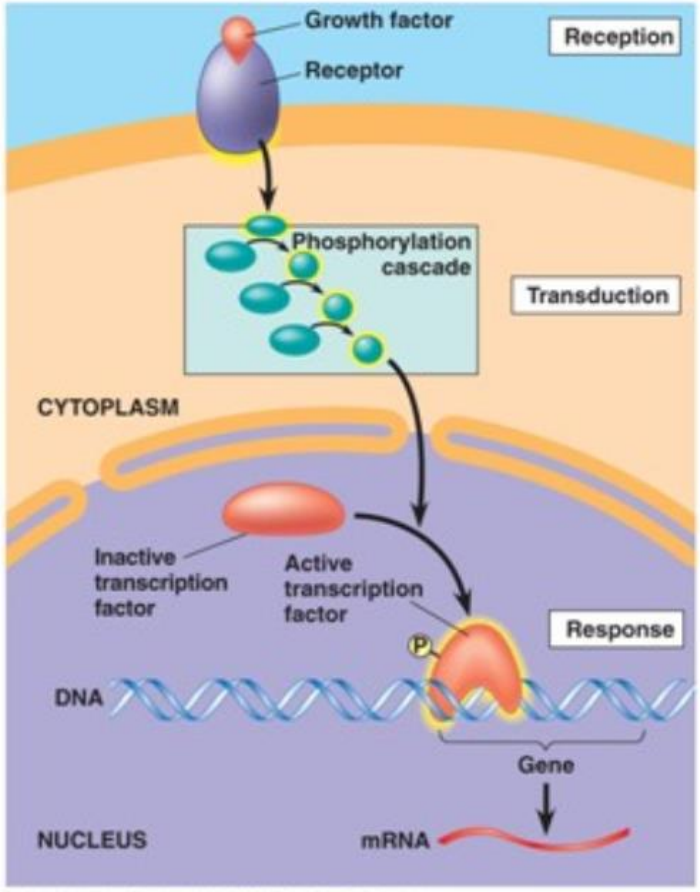
# Transmembrane proteins

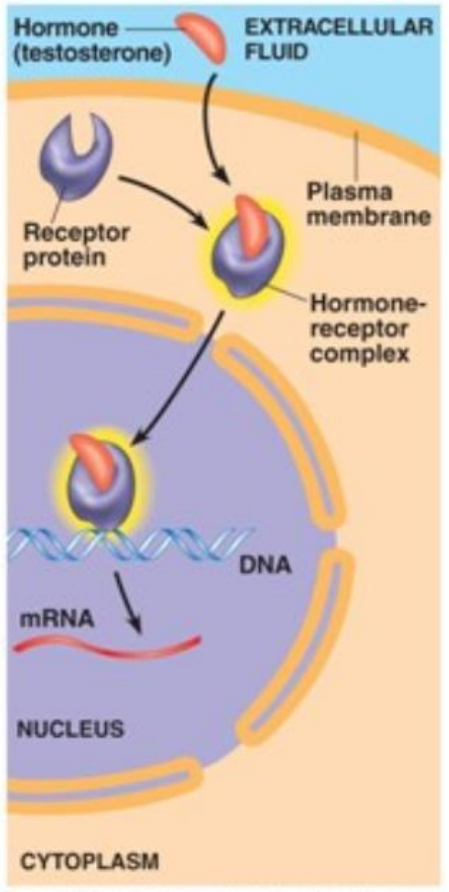
- Lipid bilayer is impermeable to most molecules
- Gates are needed – TM proteins
    - Transport
    - Signal transduction
    - Energy conversion

# Transport

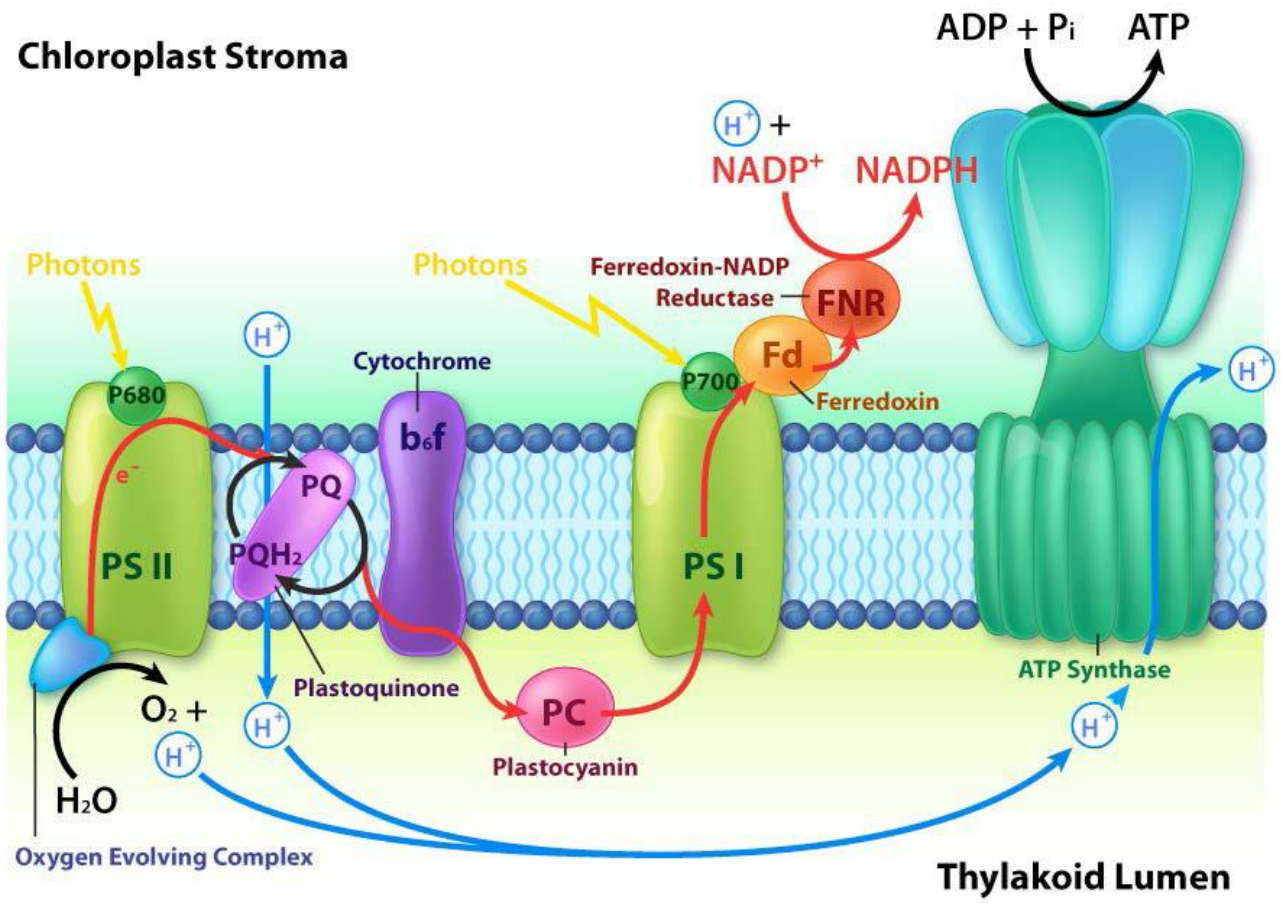

Transported molecule
Channel protein
Carrier protein
Pump
Plasma membrane
High
Energy
Low
Electrochemical potential gradient
Simple diffusion
Passive transport (in the direction of electrochemical gradient)
Primary active transport (against the direction of electrochemical gradient)

# Signal transduction



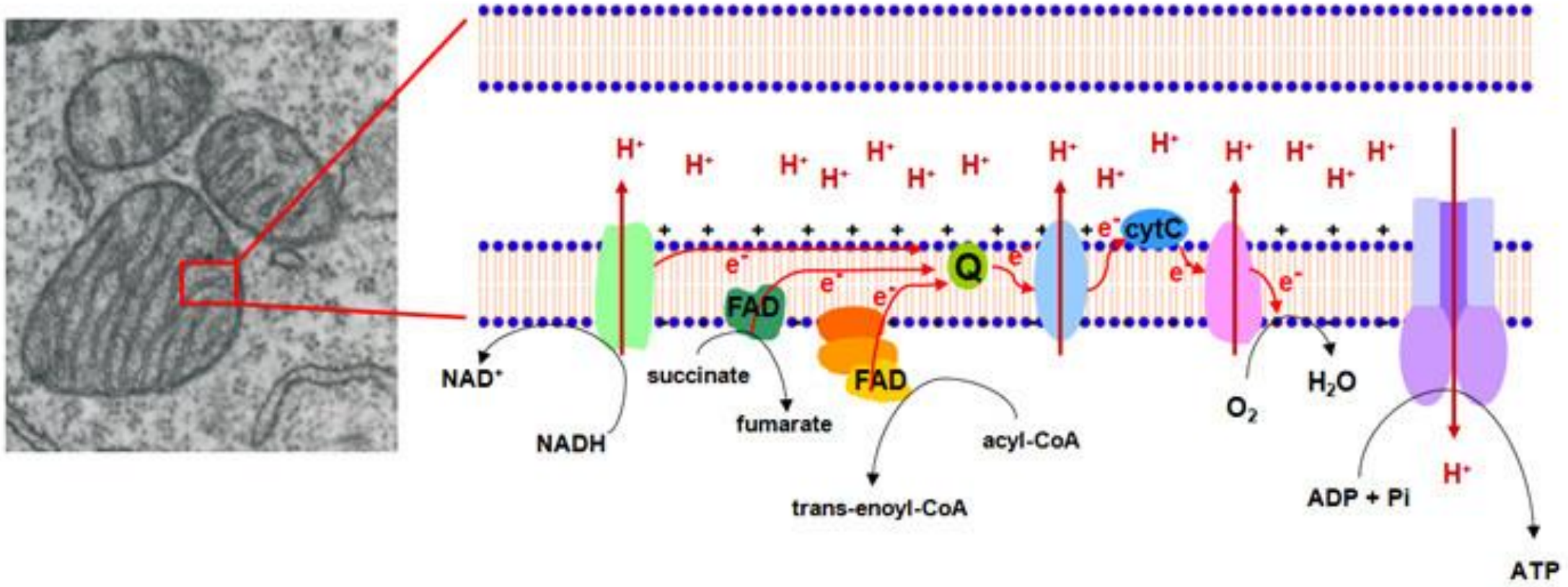



# Photosynthesis



# Mitochondrium

## Krebs cycle, aerobic respiration



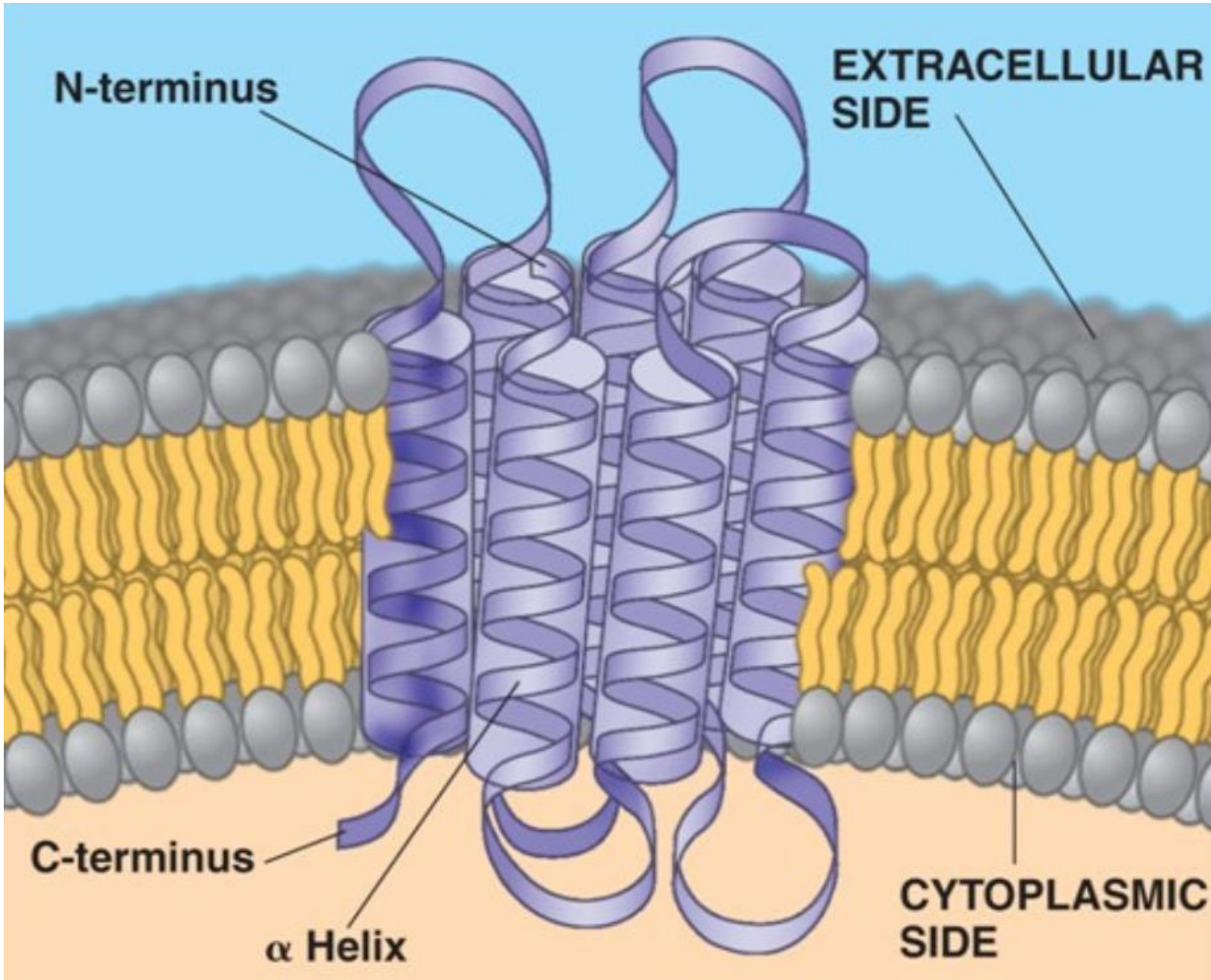
# Occurence

- Bacteria 1-2% of genome
- Eukaryotes ~25%
  - important drug targets; 134 GPCR targets
  - at least 35% of all drug targets are GPCRs
- PDB: 2%; uderrepresented; hard to crystallize

# Structure

- Alpha
- Beta

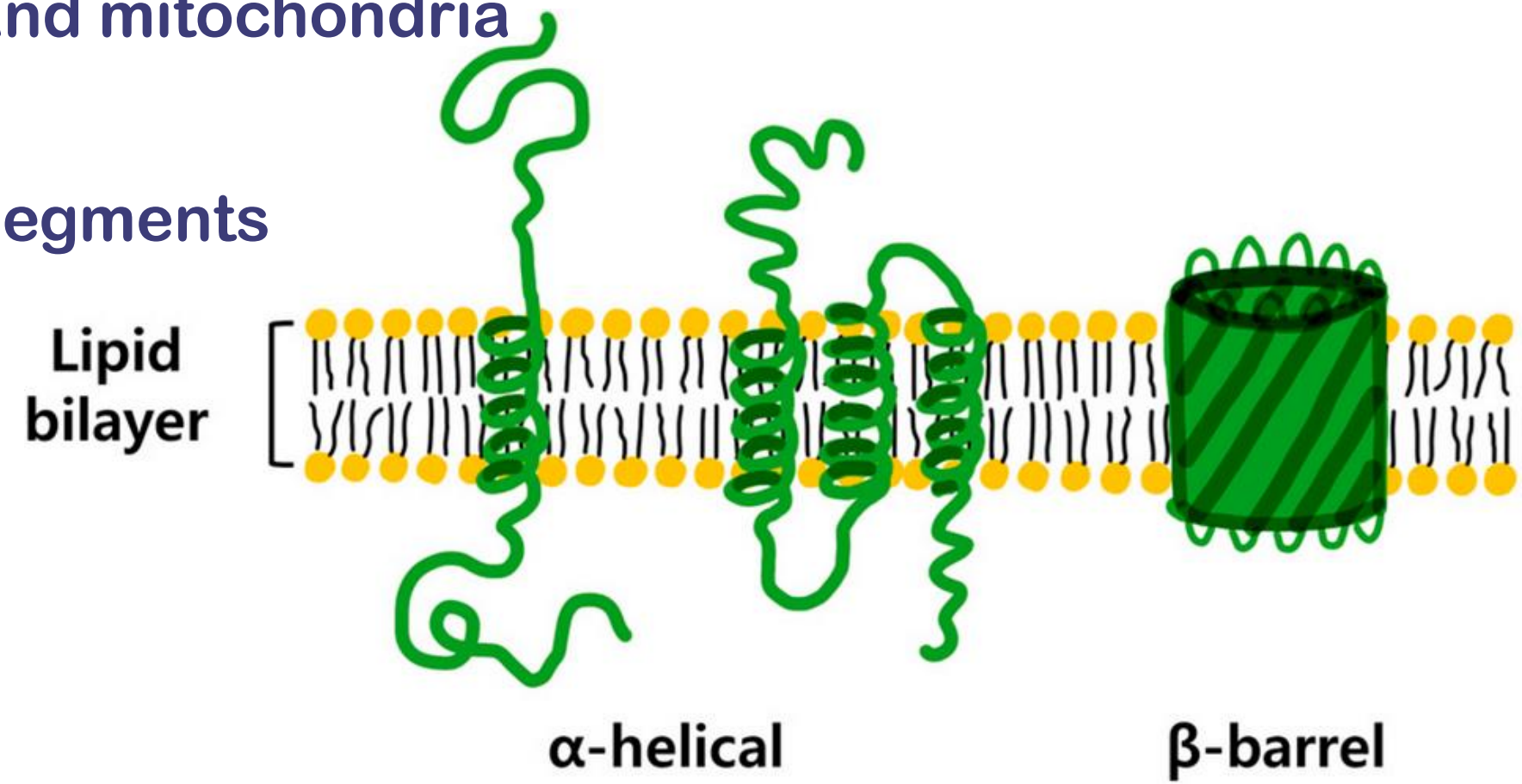
Structure is needed to understand their function



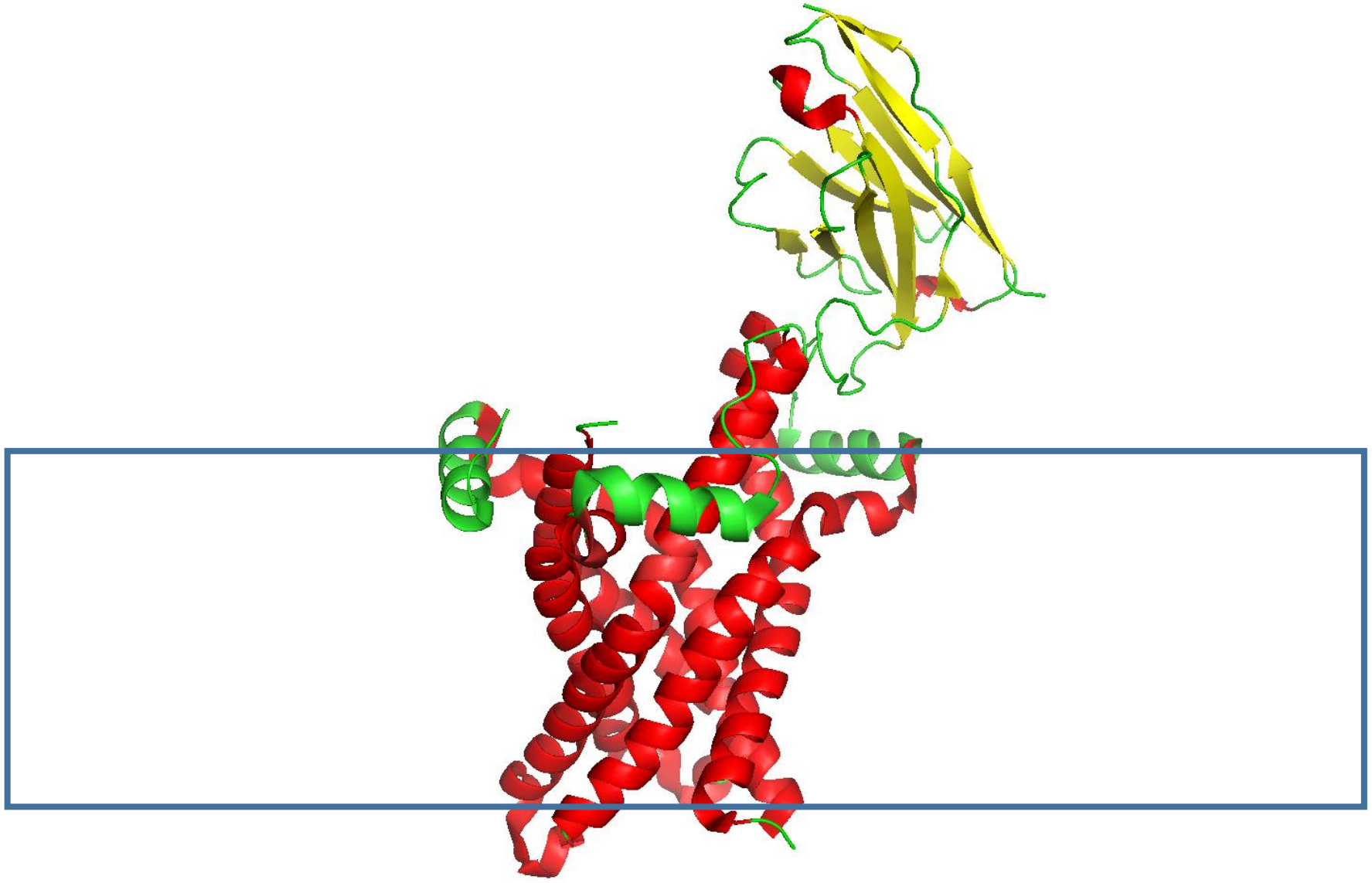
# Structure

Alpha: from unicellular to mammals
Beta: bacteria and mitochondria

Number of TM segments
1: 40%
6,7: frequent



# Orientation of TM helices

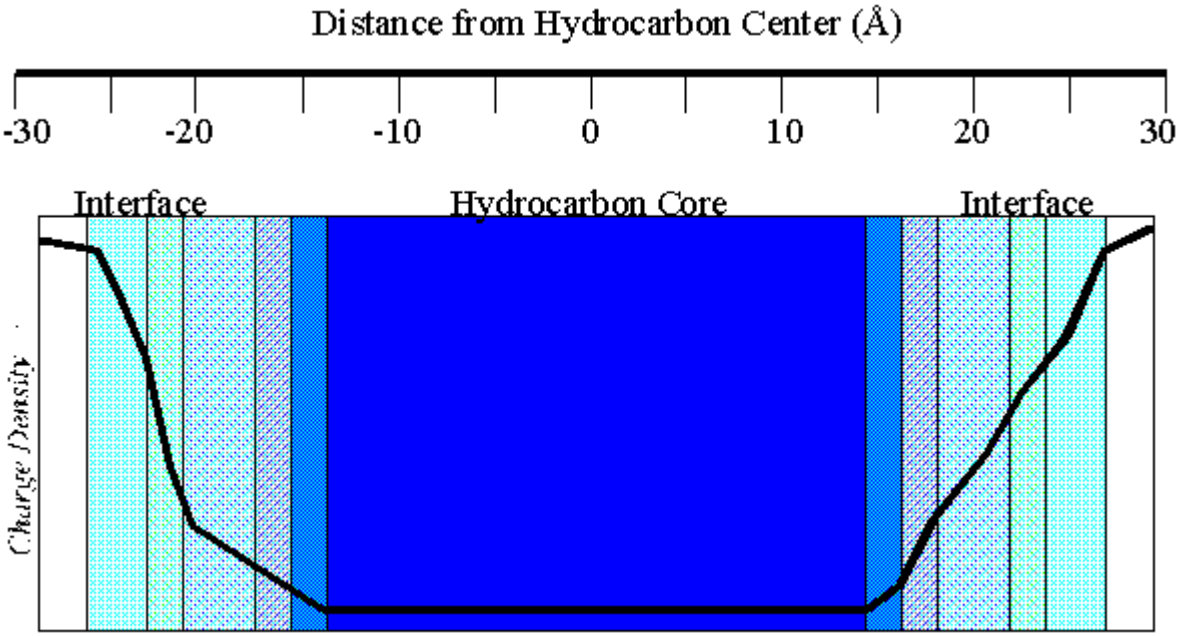


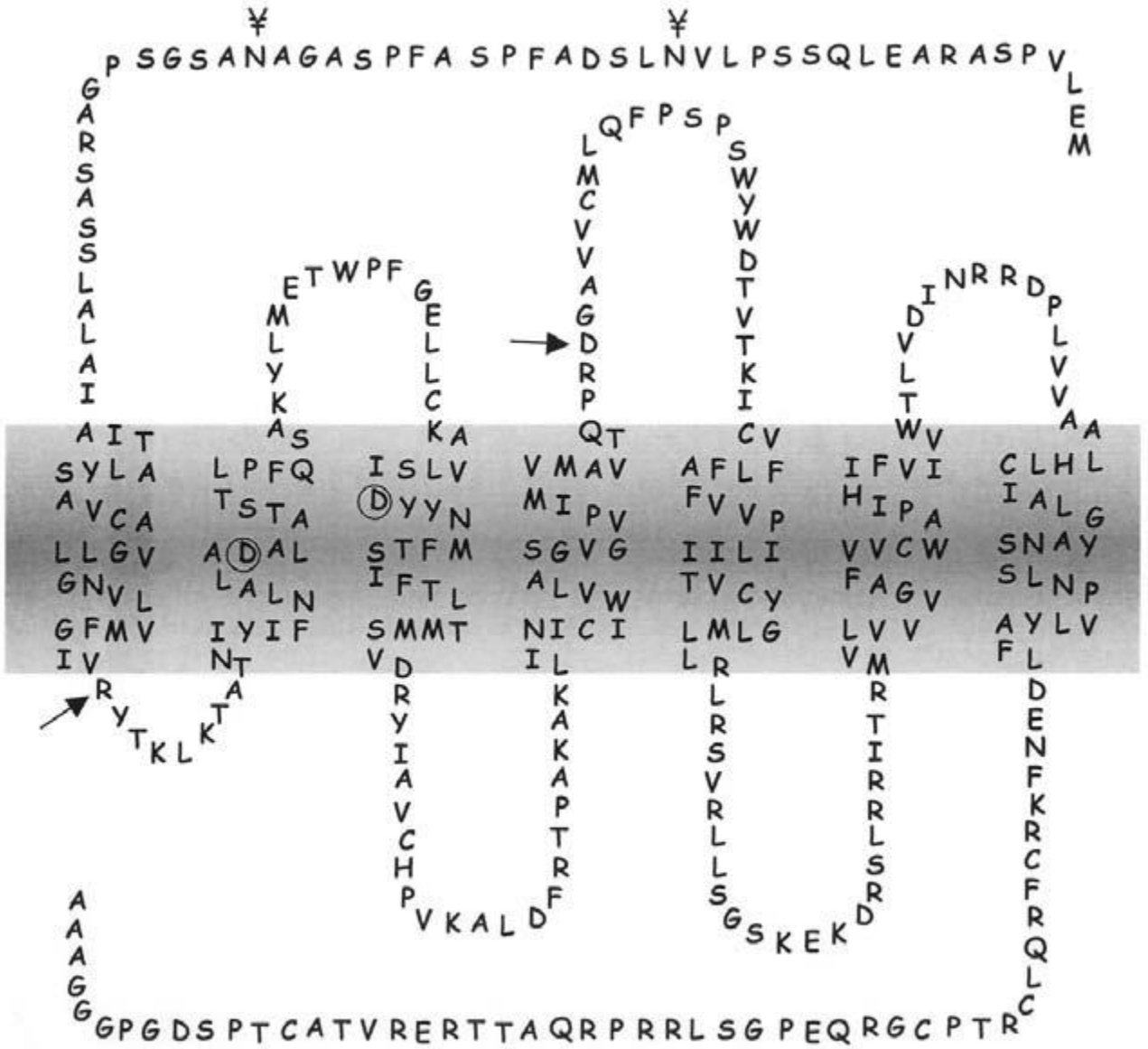
# Topography

- Position of TM regions

Size of TM regions

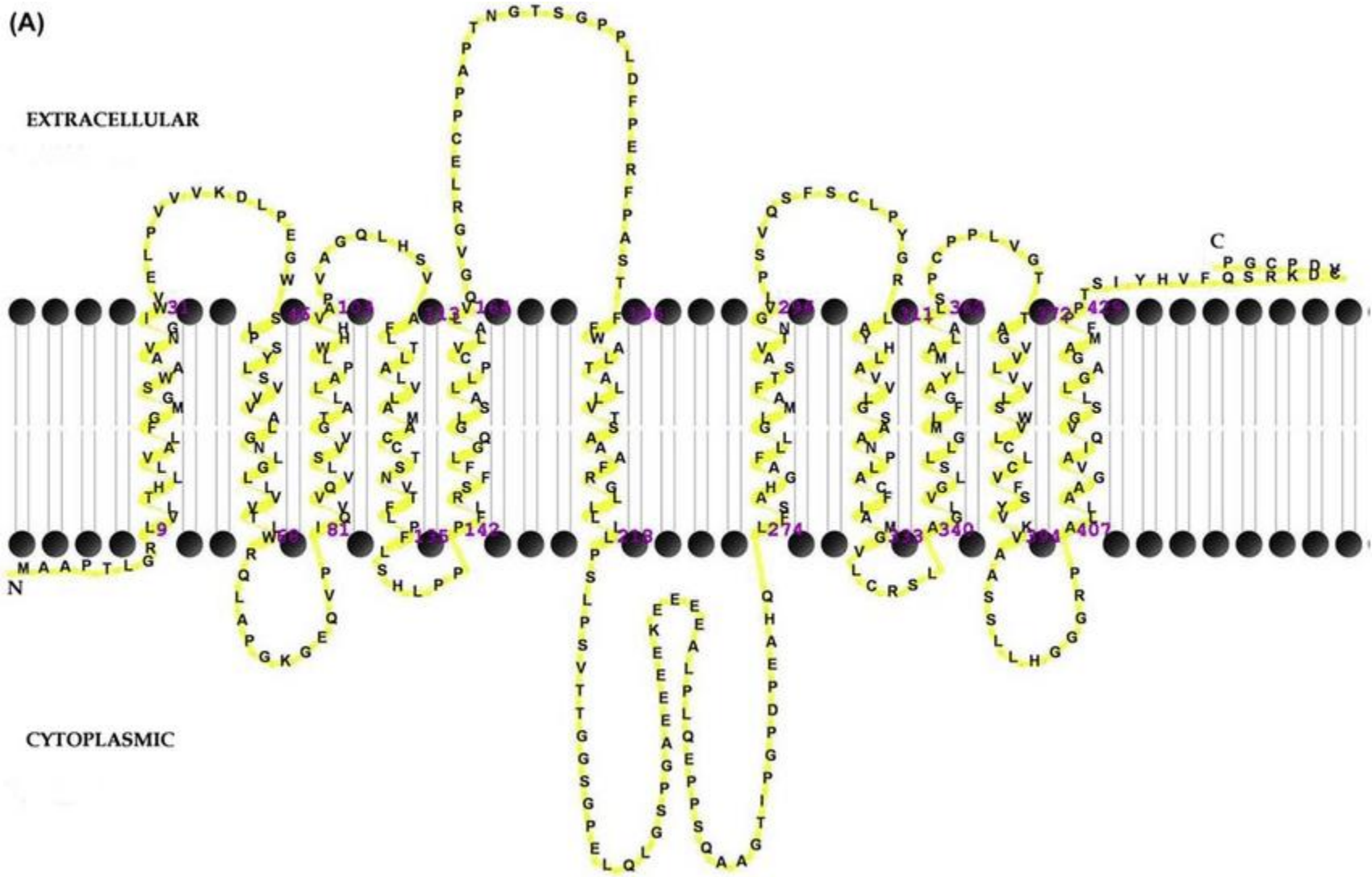
30 Å ~ 20-25 aa



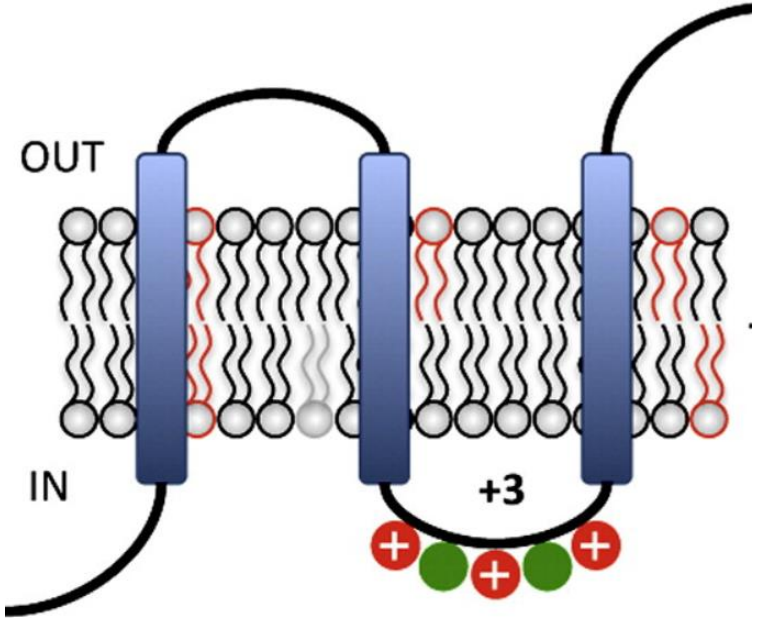



# Topology

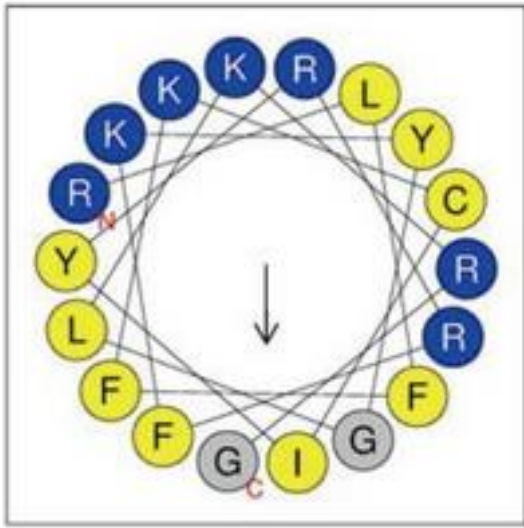
- Localization



# TM sequences




- Hydrophobic residues, low complexity
- Stucture is stabilized with internal H-bond
- Positive inside rule
- Ampiphatic helices; anchors
- Proline is a helix breaker, but present in TM helices
  - missing amide hydrogen - > partial charge
  - important for function; ion transport



# PDBTM database

## PDBTM: Protein Data Bank of Transmembrane Proteins

PDBTM version: 2019-03-15 Number of transmembrane proteins: 4115 (alpha: 3662 , beta: 429 )

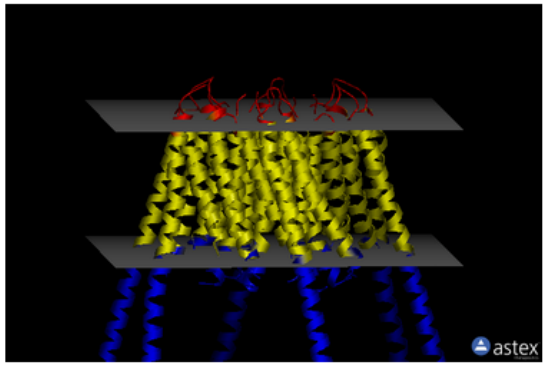
all 6m96

- Home
- Search
- Download
- Statistics
- Documents
- Help

### Welcome to the PDBTM home page

PDBTM is the first comprehensive and up-to-date transmembrane protein selection of the Protein Data Bank (PDB). PDBTM database is maintained at the Institute of Enzymology by the Membrane Protein Bioinformatics Research Group. The PDBTM database was created by scanning all PDB entries with the TMDET algorithm. You can get more information about PDBTM in our articles and in the PDBTM manual. If you find PDBTM useful in your research, please cite our articles (Bioinformatics 20, 2964-2972; Nucleic Acids Research 33 Database Issue, D275-D278; Nucleic Acids Research 41 Database Issue, D524-D529 ).

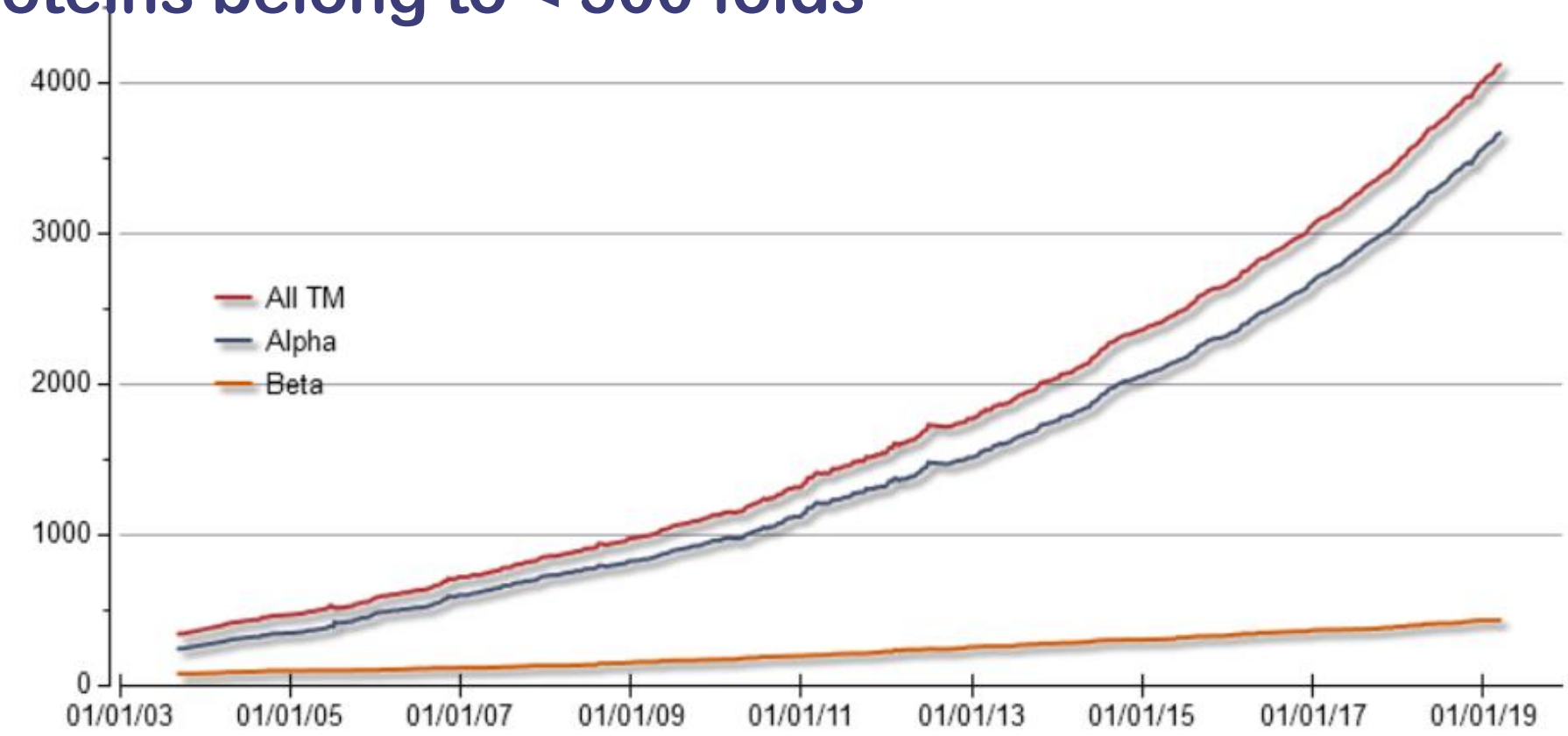
6aki



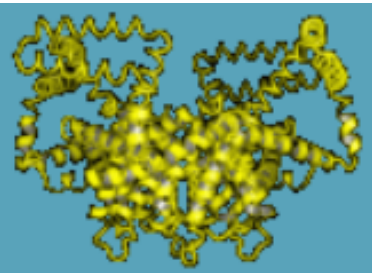

PDBTM type: Tm_Alpha
Chain(s): A[4], B[4], C[4], D[4], E[4], F[4]

# TM folds

Globular folds < 10k
Large sizes, but small variations
90% of TM proteins belong to < 500 folds



# TMDET



## Detection of transmembrane regions by using 3D structure of proteins

Version 2.0 — Mon 02 Dec, 2019

- General info
- Form
- Manual
- Faq
- Articles
- PDBTM
- Comment

### TMDET home page

TMDET may be used for the detection of the transmembrane regions of membrane proteins using their 3D structure only. You can read more about the algorithm in our articles. You can send your protein structure by using this form.

If you would like to know the membrane topology of a protein already deposited in the Protein Data Base, please use PDB_TM database.

# OPM

orientations of proteins in membranes (OPM) database

UNIVERSITY OF MICHIGAN | COLLEGE OF PHARMACY | LOMIZE GROUP

Search proteins by PDB ID or name search

HOME | ABOUT OPM | SEARCH | DOWNLOAD OPM FILES | CONTACT US | PPM SERVER | TMPFOLD SERVER

**Protein Classification**

Types (3)
Classes (11)
Superfamilies (504)
Families (976)
Species (854)
Localizations (24)
Proteins (4739)

**Assembly**

Superfamilies (9)
Families (19)
Localizations (8)
Assemblies (207)

**Protein Links**

PDB Sum, PDB, MPKS, MPDB

PPM Server

TMPfold Server

## Orientations of Proteins in Membranes (OPM) database

OPM provides spatial arrangements of membrane proteins with respect to the hydrocarbon core of the lipid bilayer.

OPM includes all unique experimental structures of transmembrane proteins and some peripheral proteins and membrane-active peptides (Features).

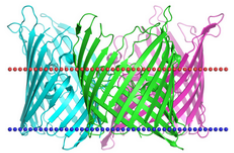
Each protein is positioned in a lipid bilayer of adjustable thickness by minimizing its transfer energy from water to the membrane (Methods).

OPM provides structural classification and sorting according to different criteria (Classification).

Our calculations are in agreement with experimental studies of 24 transmembrane and 39 peripheral peptides and proteins.

**OPM also provides a few preliminary results of our computational analysis of transmembrane α-helix association in experimental structures of selected polytopic proteins (Assembly pages).**

***For more information on single-spanning transmembrane proteins please see our Membranome database***



Porin B (PorB), different strain pdb-3wi4

Download File: 3wi4.pdb

**In citing the Orientations of Proteins in Membranes (OPM) database please refer to**
Lomize MA, Pogozheva ID, Joo H., Mosberg HI, Lomize AL (2012) OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40** (Database issue), D370-D376. PDF PubMed

**For an explanation of our method please refer to**
Lomize AL, Pogozheva ID, Lomize MA, Mosberg HI (2006) Positioning of proteins in membranes: A computational approach. *Protein Science.* **15**, 1318-1333. PDF PubMed

**For a new version of our method please refer to**
Lomize AL, Pogozheva ID, Mosberg HI (2011) Anisotropic solvent model of the lipid bilayer. 2. Energetics of insertion of small molecules, peptides, and proteins in membranes. *J Chem Inf Model.* **51**, 930-946. PDF PDF (supplementary) PubMed

**For more information on peripheral proteins please refer to**
Lomize AL, Pogozheva ID, Lomize MA, Mosberg HI (2007) The role of hydrophobic interactions in positioning of peripheral proteins in membranes. *BMC Struct Biol.* **7**, 44. PDF PubMed

**Structure Statistics (distinct protein structures/PDB entries)**

*Type:* Transmembrane - **(2579/5883)**
- *Class:* Alpha-helical polytopic - (1850/4069)
- *Class:* Bitopic proteins - (442/1308)
- *Class:* Beta-barrel transmembrane - (287/506)

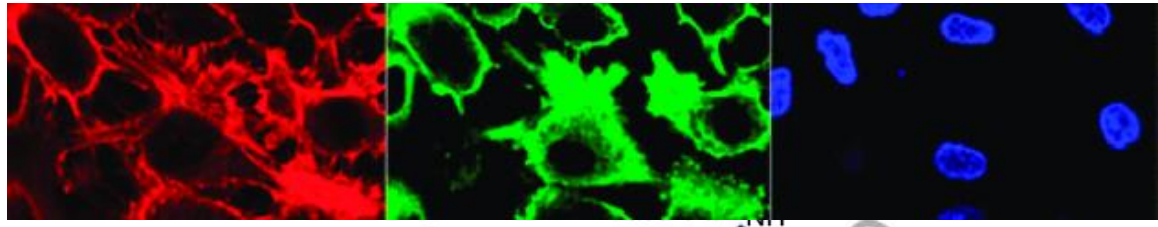
*Type:* Monotopic/peripheral - **(1513/4986)**
- *Class:* All alpha monotopic/peripheral - (320/1207)
- *Class:* All beta monotopic/peripheral - (572/1846)
- *Class:* Alpha/Beta monotopic/peripheral - (338/1304)
- *Class:* Alpha + Beta monotopic/peripheral - (283/629)

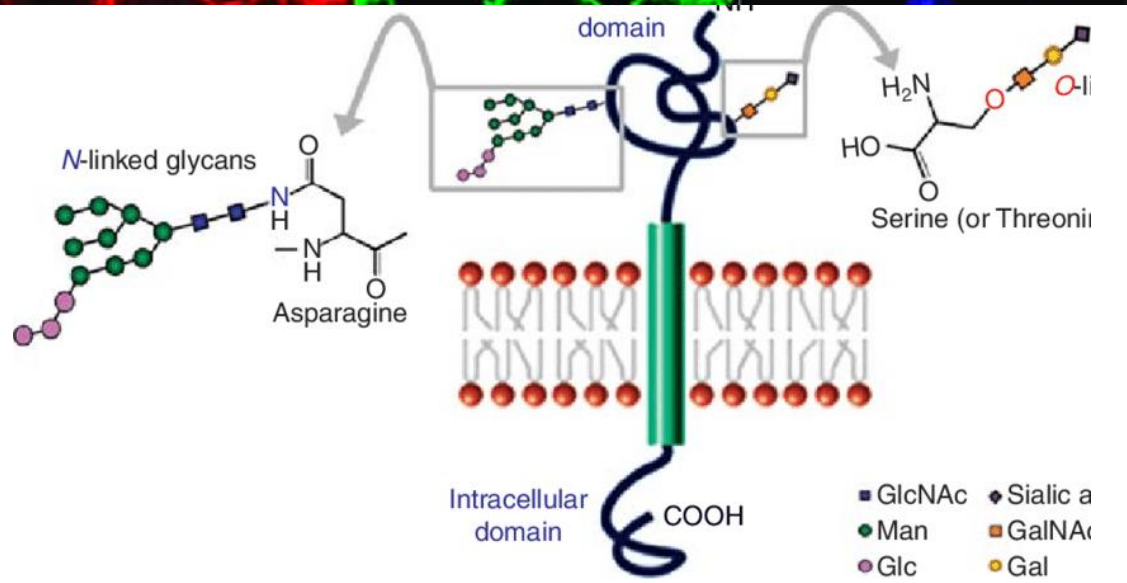
*Type:* Peptides - **(647/1159)**
- *Class:* Alpha-helical peptides - (448/787)
- *Class:* Beta-hairpin peptides - (125/229)
- *Class:* Beta-helical peptides - (17/36)
- *Class:* Peptides of nonregular structure - (57/107)

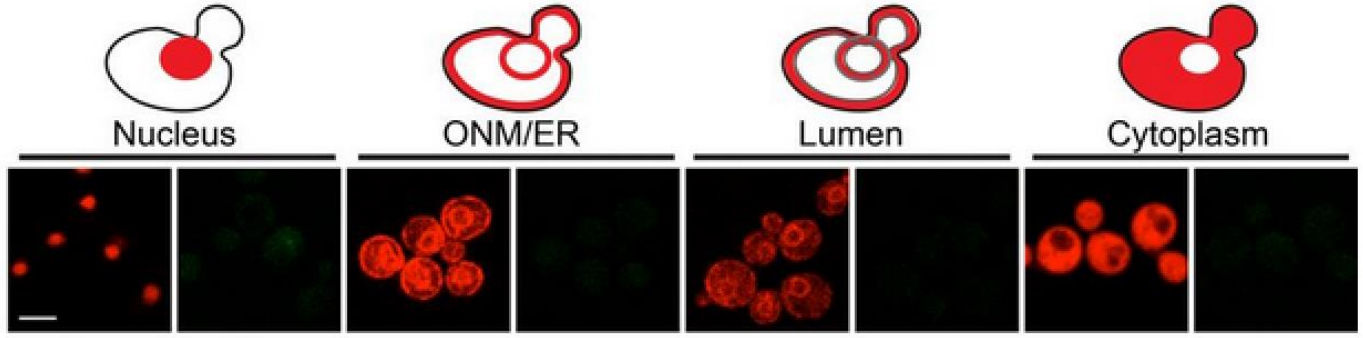
# Experimental methods

- Immuno localization; antibodies
- Glycosylation sites
- Reporter enzyme fusion; GFP

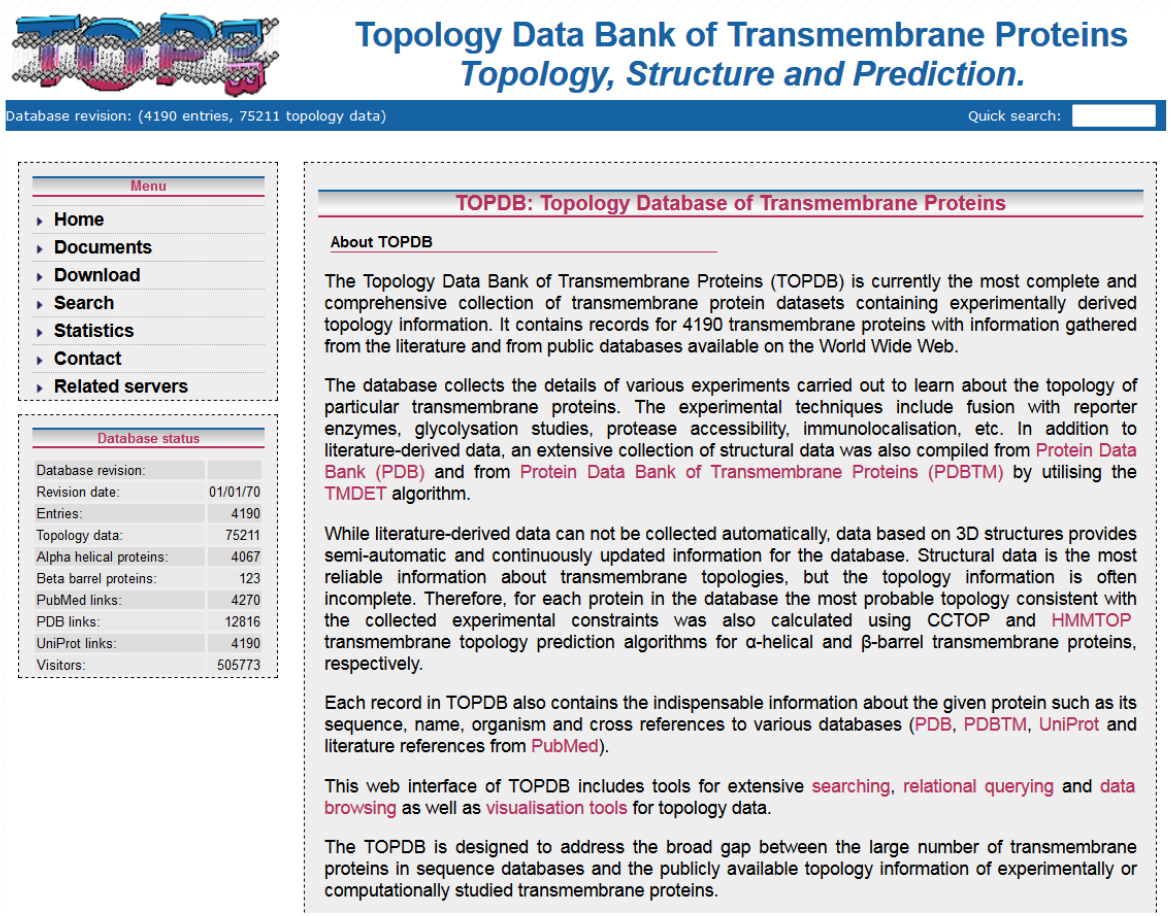
Data collected in TOPDB





# TOPDB


Topology Data Bank of Transmembrane Proteins
Topology, Structure and Prediction.
Database revision: (4190 entries, 75211 topology data)
Quick search:
Menu
Home
Documents
Download
Search
Statistics
Contact
Related servers
Database status
Database revision:
Revision date: 01/01/70
Entries: 4190
Topology data: 75211
Alpha helical proteins: 4067
Beta barrel proteins: 123
PubMed links: 4270
PDB links: 12816
UniProt links: 4190
Visitors: 505773
TOPDB: Topology Database of Transmembrane Proteins
About TOPDB
The Topology Data Bank of Transmembrane Proteins (TOPDB) is currently the most complete and comprehensive collection of transmembrane protein datasets containing experimentally derived topology information. It contains records for 4190 transmembrane proteins with information gathered from the literature and from public databases available on the World Wide Web.
The database collects the details of various experiments carried out to learn about the topology of particular transmembrane proteins. The experimental techniques include fusion with reporter enzymes, glycolysation studies, protease accessibility, immunolocalisation, etc. In addition to literature-derived data, an extensive collection of structural data was also compiled from Protein Data Bank (PDB) and from Protein Data Bank of Transmembrane Proteins (PDBTM) by utilising the TMDET algorithm.
While literature-derived data can not be collected automatically, data based on 3D structures provides semi-automatic and continuously updated information for the database. Structural data is the most reliable information about transmembrane topologies, but the topology information is often incomplete. Therefore, for each protein in the database the most probable topology consistent with the collected experimental constraints was also calculated using CCTOP and HMMTOP transmembrane topology prediction algorithms for α-helical and β-barrel transmembrane proteins, respectively.
Each record in TOPDB also contains the indispensable information about the given protein such as its sequence, name, organism and cross references to various databases (PDB, PDBTM, UniProt and literature references from PubMed).
This web interface of TOPDB includes tools for extensive searching, relational querying and data browsing as well as visualisation tools for topology data.
The TOPDB is designed to address the broad gap between the large number of transmembrane proteins in sequence databases and the publicly available topology information of experimentally or computationally studied transmembrane proteins.

# Methods use for TM proteins

- Signal peptide identification
- Differentiate TM and globular sequences (DAS-TMfilter)
- Topology prediction (HMMTOP)
- Fold recognition (TOPDOM)
- Homology modelling
- Membrane positioning (TMDET, OPM)

# TOPDOM


## TOPDOM: Conservatively Located Domains and Motifs in Proteins

Home | Documents | Download | Search | Related sites

### TOPDOM is

a collection of domains and sequence motifs located conservatively in one side of membranes either in transmembrane or globular proteins. The database was created by predicting the transmembrane status and topology of all protein sequence in **UniProt** database by the **CCTOP** algorithm and scanning by specific domain or motif detecting algorithms. The identified domain or motif was added to the database if it was uniformly annotated in the same side of the membrane of the various proteins in **UniProt** database. The sequences in the **UniProt** database were scanned by using hmmpfam algoritm with **Pfam** and **Smart** domain databases, prosite scan with **Prosite** motif database and fingerprints algorithm with **Prints** database. For further details refer to our **manuscripts**.

### TOPDOM can be

**downloaded** either the whole database or various subset of the database in XML format. The database can be searched by **keywords** or **identifiers**. Moreover, users can **scan** their protein sequences against the TOPDOM domains and motifs.

### Constrained prediction

can be performed after sequence **scanning** by the **CCTOP** transmembrane topology prediction algorithm using the search results as constrains.

**Statistics**

| | |
|---|---|
| Visitors | 444186 |
| Entries | 6232 |
| Pfam | 3930 |
| Prints | 589 |
| Prosite | 1375 |
| Smart | 338 |
| UniProt | 195371 |
| Inside | 4440 |
| Outside | 1792 |
| Transmem | 1331 |
| Search by sq | 1109268 |
| Search by kw | 5679 |
| Search by id | 12422 |

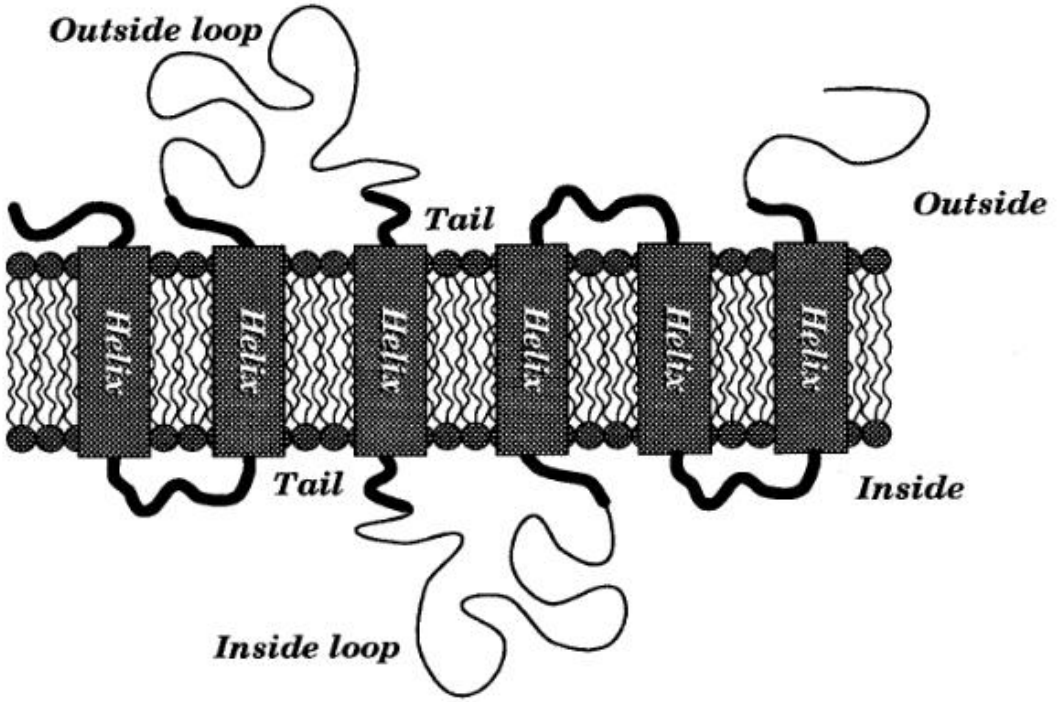
# Topology prediction

- Structure determination if difficult
- Prediction methods are needed
  - Machine learning
    - Neural networks
    - Hidden Markov model
    - Support Vector Machine

# HMMTOP

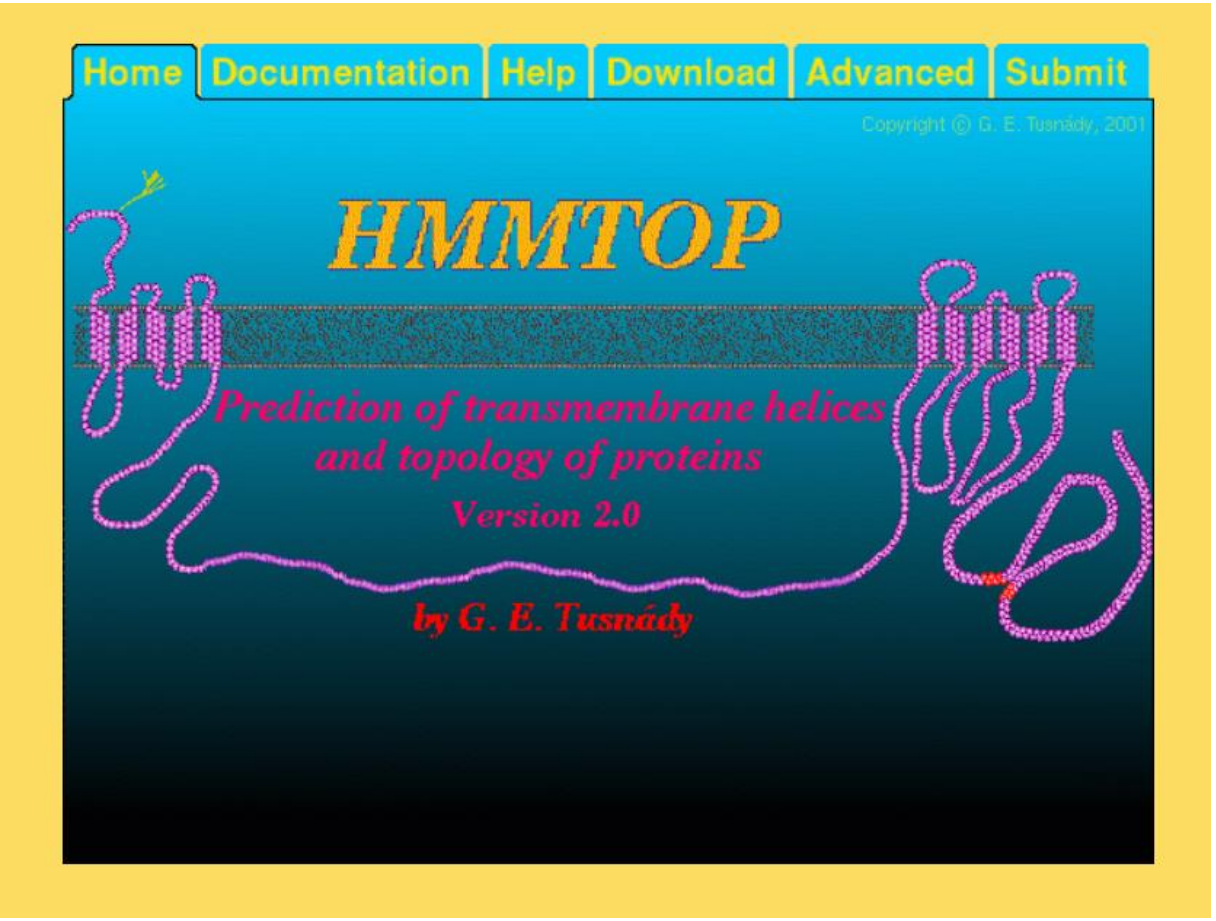
Hypothesis: the topology is determined by the difference in the amino acid distributions.

o->h->i->h->o hidden Markov model

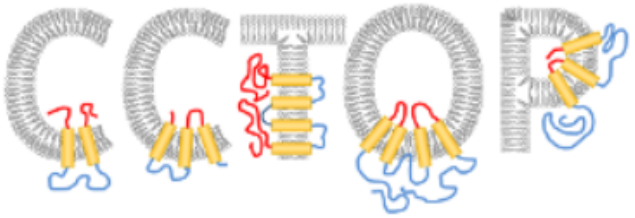



The most likely topology is the one that maximizes this difference.

# Incorporation of experimental data




# CCTOP



## Constrained Consensus TOPology prediction server

Documents

Predictions

| | |
|---|---|
| Visitors: | 1206895 |
| Prediction count: | 67769 |
| Server version: | s.1.00 |
| Last changed: | 13.07.2017 |

### Welcome to the Constrained Consensus TOPology prediction server.

Given the amino acid sequence of a putative α-helical transmembrane protein, CCTOP predicts its topology i.e. localization of membrane spanning regions and orientation of segments between them. The prediction is a consensus of 10 different methods enchanced with available structural and experimental information of any homologous proteins in the TOPDB database. CCTOP was tested on a benchmark set containing 170 proteins with known structure and achieved the highest accuracy among state-of-art and consensus methods.

You can find more information how CCTOP works in the description or in the article. A short description about the server can be found in the manual.

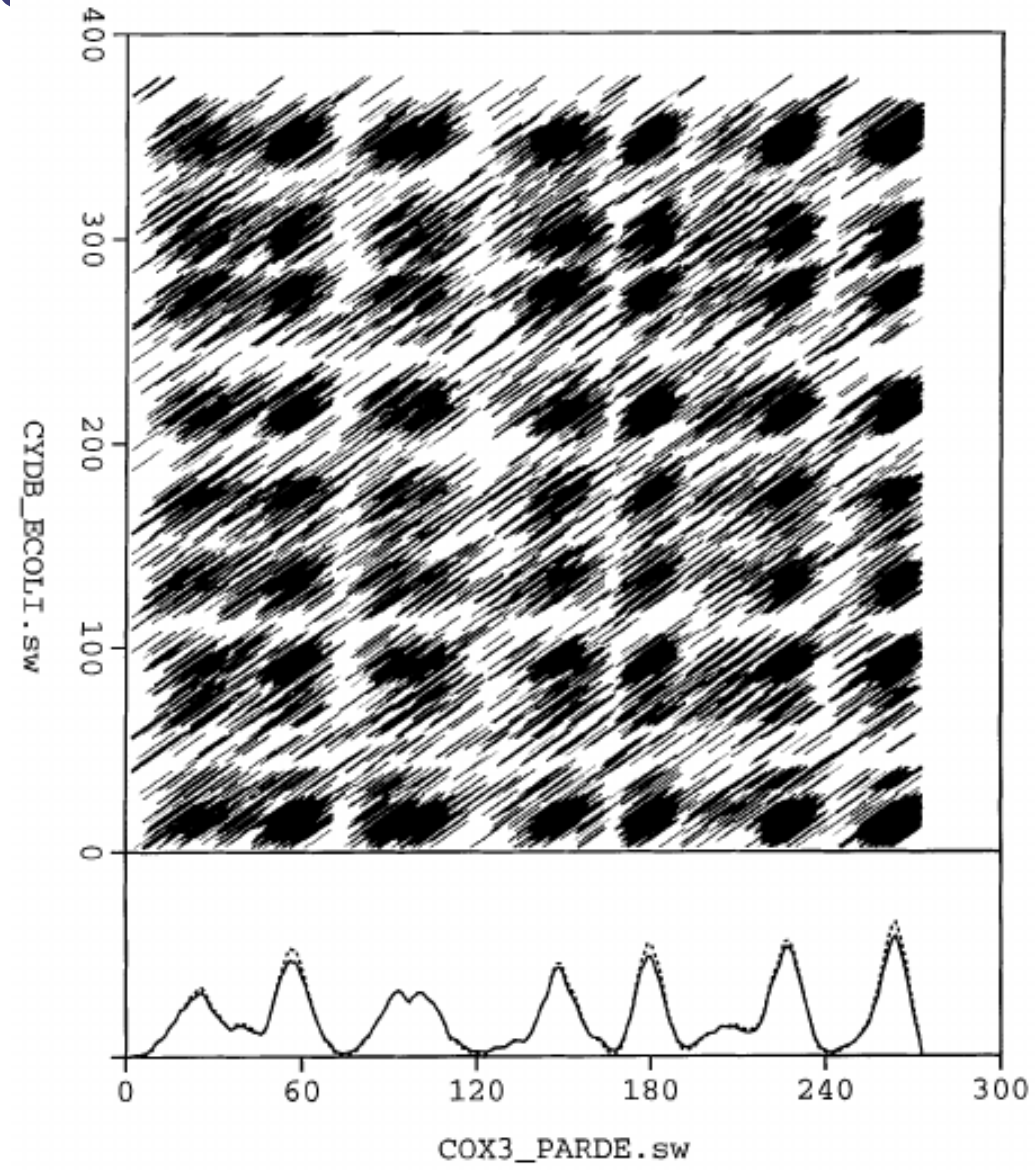
You can submit protein sequences here.

# Other methods

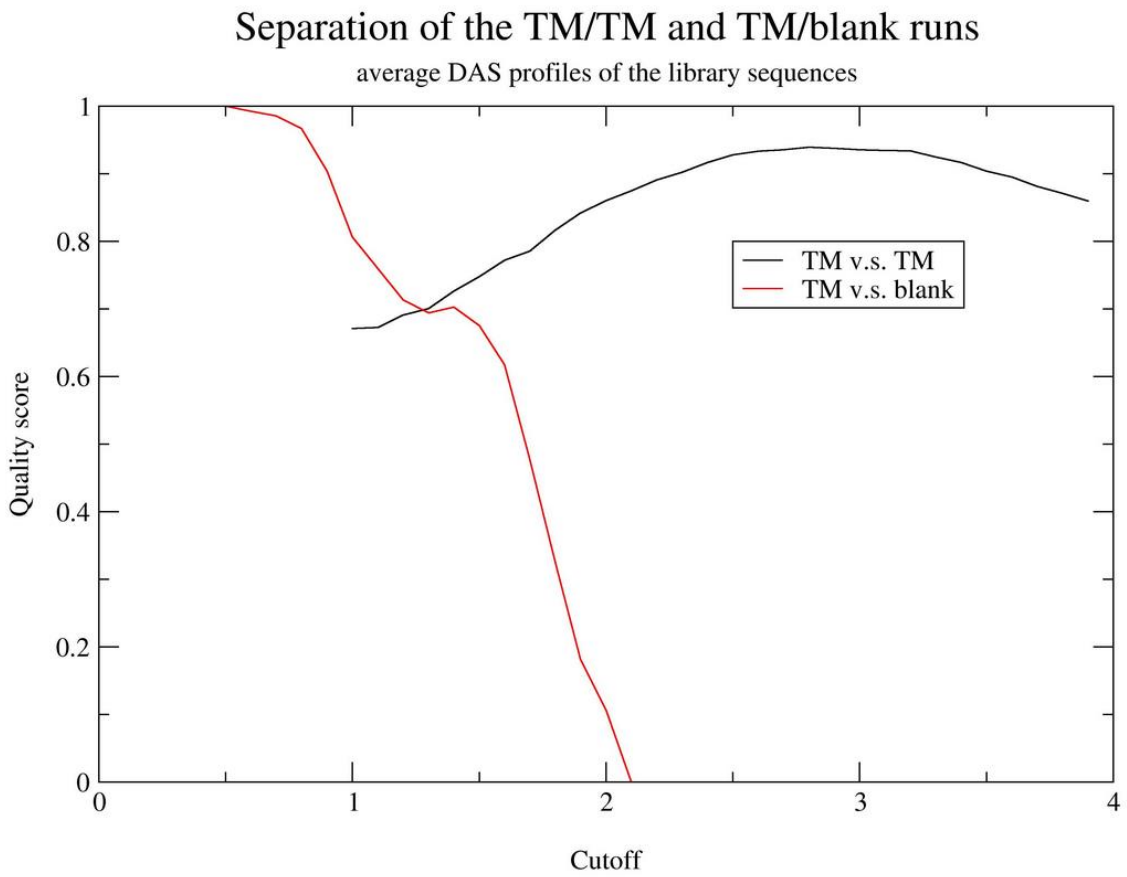
- DAS - Prediction of transmembrane regions in prokaryotes using the Dense Alignment Surface method (Stockholm University)
- MemBrain - Transmembrane protein structure prediction (HongBin Lab)
- Memsat - Membrane Helix Prediction (University College London)
- Octopus - Prediction of membrane protein topology and signal peptides (Stockholm University)
- Philius - Transmembrane Topology Prediction Server (Yeast Resource Center)
- Phobius - A combined transmembrane topology and signal peptide predictor (University of Copenhagen)
- PredictProtein - Prediction of transmembrane helix location and topology (Columbia University)
- Pro and Prodiv - Alpha-helical transmembrane protein topology prediction methods utilizing hidden Markov models and evolutionary information (Stockholm University)
- SignalP - Predicts the presence and location of signal peptide cleavage sites (Technical University of Denmark)
- Scampi - Prediction of membrane protein topology from first principles (Stockholm University)
- SOSUI - Prediction of transmembrane regions (Nagoya University, Japan)
- TMHMM - Prediction of transmembrane helices in proteins (CBS; Denmark)
- TMpred - Prediction of transmembrane regions and protein orientation (EMBnet-CH)
- TopPred - Topology prediction of membrane proteins (France)

# Dense Alignment Surface method (DAS)

Hydrophobic TM helices are similar to each other

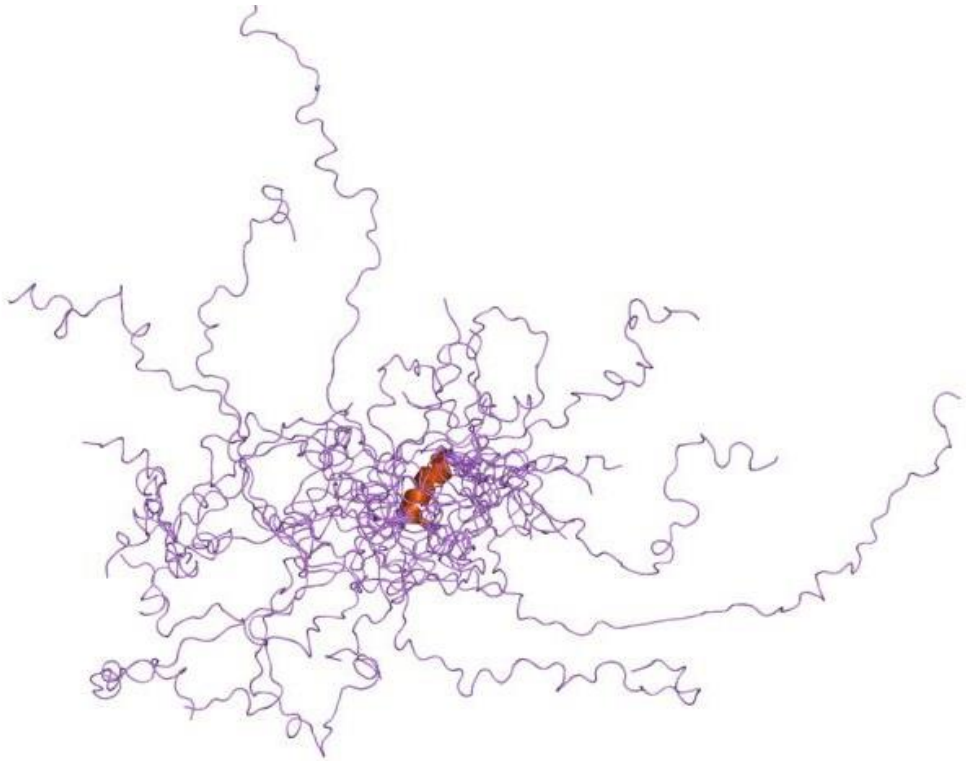


# DAS-TMfilter



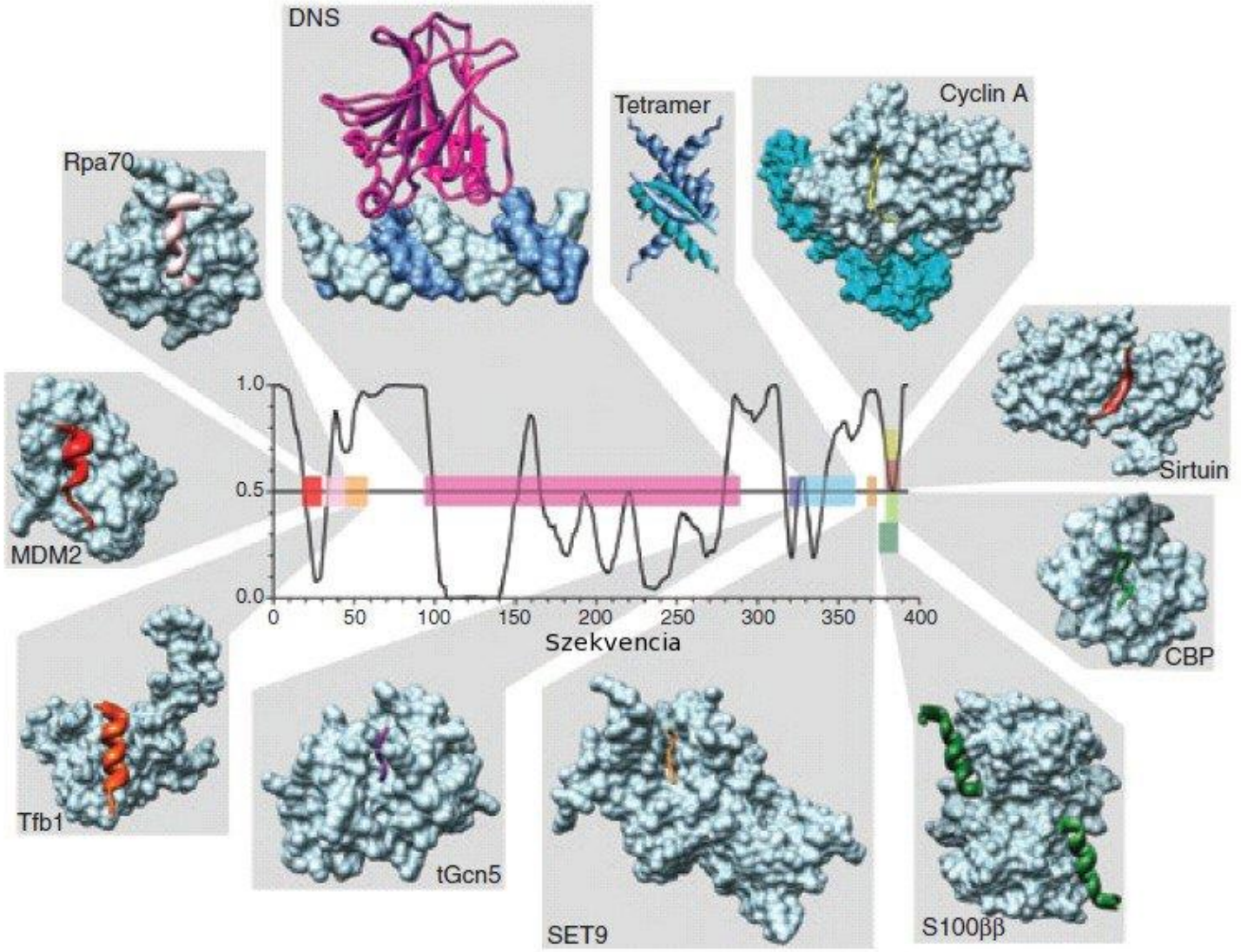
# Intrinsically Disordered Proteins

- Exception to the Anfinsen dogma
- IDP ~ IUP
- lack of stable (steady) 3D structure
- highly dynamics structures
- usually many interaction partners
- different amino acid composition
  usually with low complexity



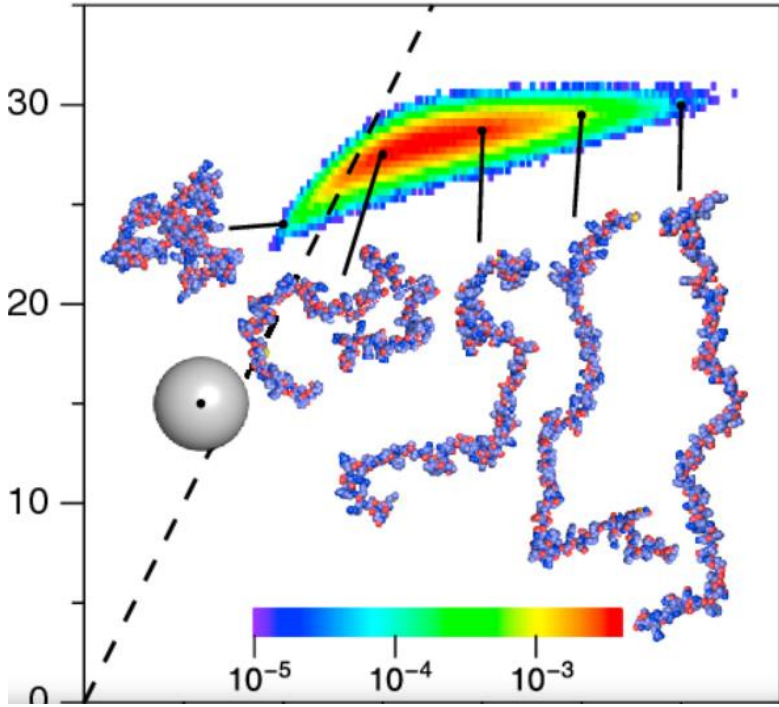
# Cellular tumor antigene p53

regulator of the cell cycle and apoptosis



# Experimental validation

- Lack of secondary structure; CD spectroscopy
- Lack of X-ray electron density
- Radius of gyration



# DisProt database

DisProt Browse Release notes Download Help About Search DisProt... Search Login Feedback

DisProt

Version: **8.0**
Release: **2019_09**

## Intrinsically disordered proteins

DisProt is a database of intrinsically disordered proteins. Disordered regions are manually curated from literature. DisProt annotations cover both structural and functional aspects of disorder detected by specific experimental methods. Annotation concepts and detection methods are encoded in the Disorder Ontology. Read more about DisProt

Search in DisProt ... Search

Example 1 Example 2

### Proteins per organism

*H. sapiens:* 578
*M. musculus:* 88
*R. norvegicus:* 50
*S. cerevisiae:* 128
Viruses: 126
*E. coli:* 7
*A. thaliana:* 33
*D. melanogaster:* 30
*C. elegans:* 13
Fungi: 156

### Statistics

| | Total | Not ambiguous |
|---|---|---|
| **Proteins** | 1.6k | 1.4k |
| **Regions** | 3.5k | 3k |
| **Residues** | 164.1k | 141.4k |
| **Disorder content** | 19.7% | 18.7% |

### Info

How to cite

Hatos A et al. DisProt: intrinsic protein disorder annotation in 2020
Nucleic Acids Res., 2019. [NAR] [PubMed]

Piovesan D et al. DisProt 7.0: a major update of the database of disordered proteins
Nucleic Acids Res., 2016. [NAR] [PubMed]

API

REST API documentation here

### Integrated resources

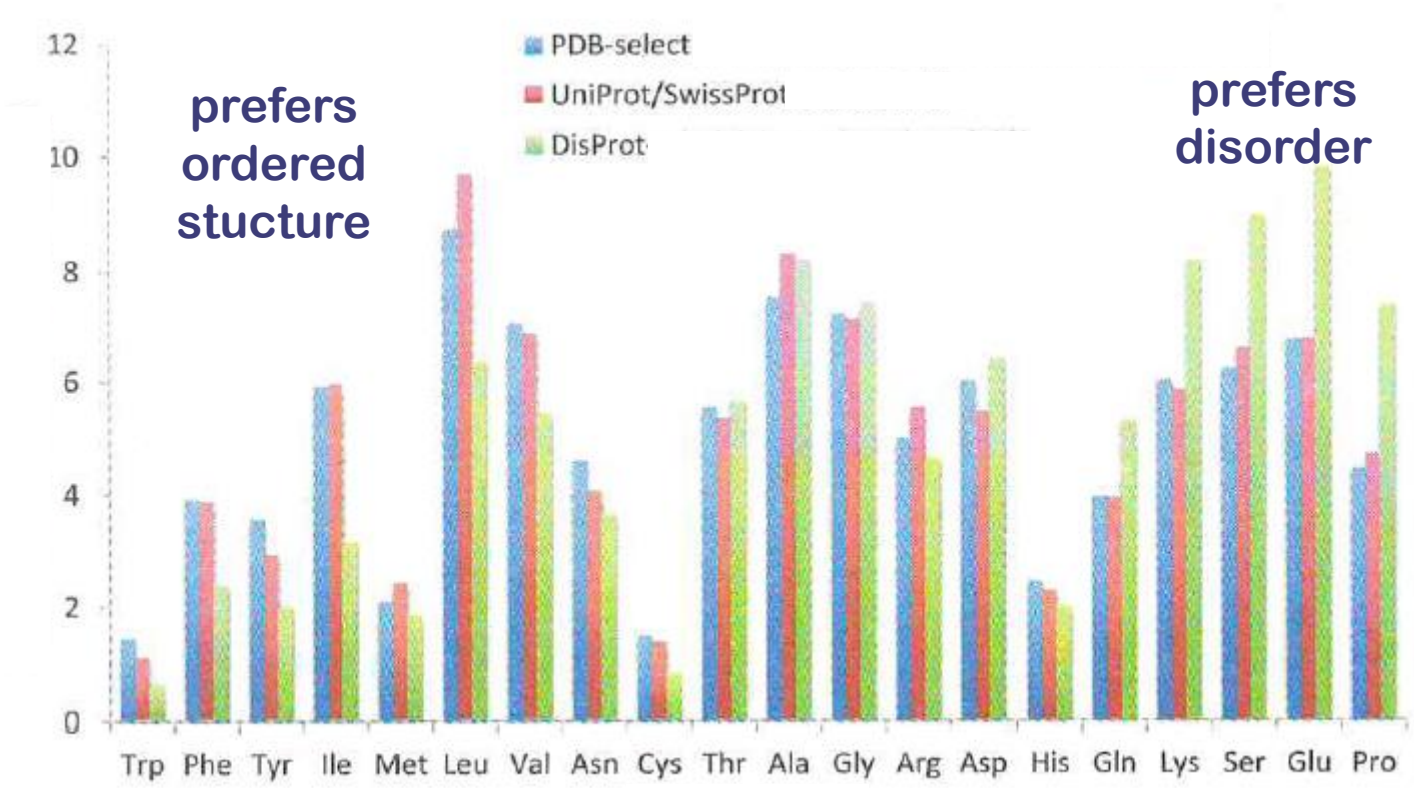
UniProt CATH Pfam Europe PMC BITEM

# Order / disorder promoting residues

| Scale | Order-promoting amino acid residues | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | W | F | Y | I | M | L | V | N | C | T |
| Top-IDP | -0.884 | -0.697 | -0.510 | -0.486 | -0.397 | -0.326 | -0.121 | 0.007 | 0.02 | 0.059 |
| B-value | 0.938 | 0.934 | 0.981 | 0.977 | 0.963 | 0.982 | 0.968 | 1.022 | 0.939 | 0.998 |
| FoldUnfold | 28.48 | 27.18 | 25.93 | 25.71 | 24.82 | 25.36 | 23.93 | 18.49 | 23.52 | 19.81 |
| DisProt | -0.465 | -0.381 | -0.427 | -0393 | 0.197 | -0.260 | -0.302 | -0.106 | -0.546 | -0.116 |

| Scale | Disorder-promoting amino acid residues | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | G | R | D | H | Q | K | S | E | P |
| Top-IDP | 0.06 | 0.166 | 0.180 | 0.192 | 0.303 | 0.318 | 0.586 | 0.341 | 0.736 | 0.987 |
| B-value | 0.994 | 1.018 | 1.026 | 1.022 | 0.967 | 1.041 | 1.029 | 1.025 | 1.052 | 1.050 |
| FoldUnfold | 19.89 | 17.11 | 21.03 | 17.41 | 21.72 | 19.23 | 18.19 | 17.67 | 17.46 | 17.43 |
| DisProt | 0.042 | 0.095 | 0.211 | 0.127 | -0.127 | 0.381 | 0.370 | 0.201 | 0.469 | 0.419 |

# Residue frequency




unfoldability by TOP-IDP scale ⟶ increasing

# Dentin sialophosphoprotein

## Calcium binding, mineralizatrion of dentin

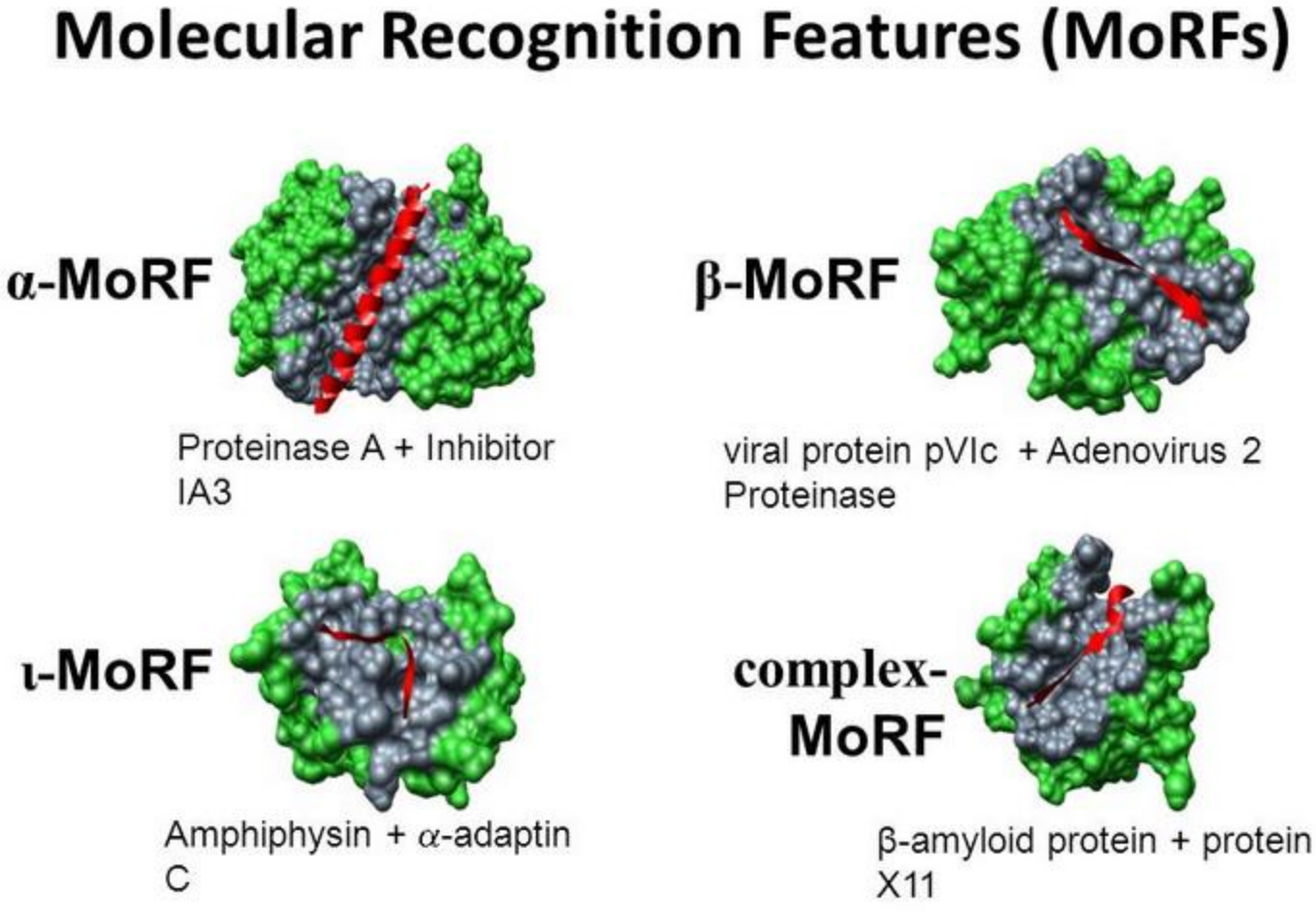
>sp|Q9NZW4|DSPP_HUMAN Dentin sialophosphoprotein OS=Homo sapiens OX=9606 GN=DSPP PE=1 SV=2
MKIITYFCIWAVAWAIPVPQSKPLERHVEKSMNLHLLARSNVSVQDELNASGTIKESGVLVHEGDRGRQENTQDGHKGEGNGSKWAEVGGKSFSTYSTLANEEGNIEGWNGDTGKAETYG
HDGIHGKEENITANGIQGQVSIIDNAGATNRSNTNGNTDKNTQNGDVGDAGHNEDVAVVQEDGPQVAGSNNSTDNEDEIIENSCRNEGNTSEITPQINSKRNGTKEAEVTPGTGEDAGLD
NSDGSPSGNGADEDEDEGSGDDEDEEAGNGKDSSNNSKGQEGQDHGKEDDHDSSIGQNSDSKEYYDPEGKEDPHNEVDGDKTSKSEENSAGIPEDNGSQRIEDTQKLNHRESKRVENRIT
KESETHAVGKSQDKGIEIKGPSSGNRNITKEVGKGNEGKEDKGQHGMILGKGNVKTQGEVVNIEGPGQKSEPGNKVGHSNTGSDSNSDGYDSYDFDDKSMQGDDPNSSDESNGNDDANSE
SDNNSSSRGDASYNSDESKDNGNGSDSKGAEDDDSDSTSDTNNSDSNGNGNNGNDDNDKSDSGKGKSDSSDSDSSDSSNSSDSSDSDSSDSDSSDSNSSSDSDSSDSDSSDSDSSDSDSSDSSN
SSDSSDSSDSSDSSDSSDSSDSKSDSSKSESDSSDSDSKSDSSDSNSSDSSDNSDSSDSSNSSNSSDSSDSSDSSDSSSSSDSSNSSDSSDSSDSSNSSESSDSSDSSDSDSSDSSDSSN
SNSSDSDSSNSSDSSDSSNSSDSSDSSDSSNSSDSSDSSDSSNSSDSSDSSDSSDSSDSSNSSDSNDSSNSSDSSDSSNSSDSSNSSDSSDSSDSDSDSSNSSDSSNSSDSSDSSNSSD
SSDSSDSSDGSDSDSSNRSDSSNSSDSSDSSDSSNSSDSSDSSDSNESSNSSDSSDSSNSSDSDSSDSSNSSDSSDSSNSSDSSESSNSSDNSNSSDSSNSSDSSDSSDSSNSSDSSNSS
DSSNSSDSSDSNSSDSSDSSNSSDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSSNSSDSSNSSDSSNSSDSSDSSDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSSDSSDSSDSSD
SSESSDSSDSSNSSDSSDSSDSSDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSSDSSNSSDSSDSSESSDSSDSSDSSDSSDSSDSSDSSDSSDSSNSSDSSDSSDSSDSSDSSDSSD
SSDSSDSSDSSDSSDSSDSSDSSDSSDSSDSNESSDSSDSSDSSDSSNSSDSSDSSDSSDSTSDSNDESDSQSKSGNGNNNGSDSDSDSEGSDSNHSTSDD

# SLiMs: Short Linear Motifs

## protein sequences mediating protein-protein interactions

linear, no real 3D structure

MoRFs
Molecular Recognition Features
motif-like peptides in disordered regions
disordered -> ordered

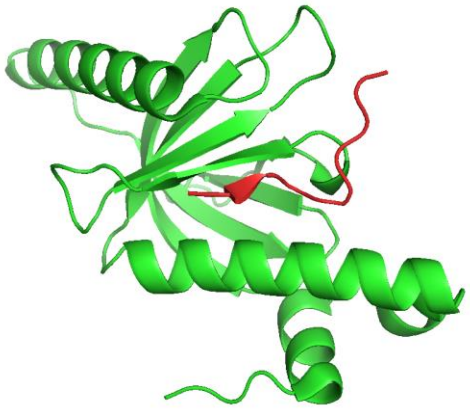


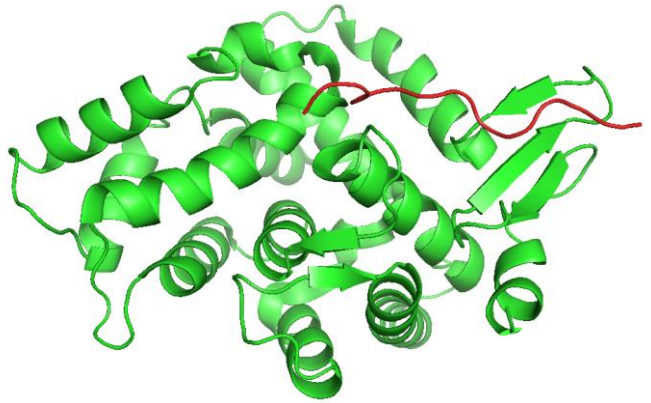
# ELMs

## Eukaryotic Linear Motif resource

| Accession | Acc. Gene-, Name | Start | End | Subsequence | Logic |
|---|---|---|---|---|---|
| ELMI001563 | Q14114 LRP8 LRP8_HUMAN | 860 | 867 | KNTKSMNFDNPVYRKTTEEE | TP |

| Accession | Acc. Gene-, Name | Start | End | Subsequence | Logic |
|---|---|---|---|---|---|
| ELMI001104 | Q14118 DAG1 DAG1_HUMAN | 889 | 892 | KGSRPKNMTPYRSPPPYVPP | TP |





# Prediction of disorder

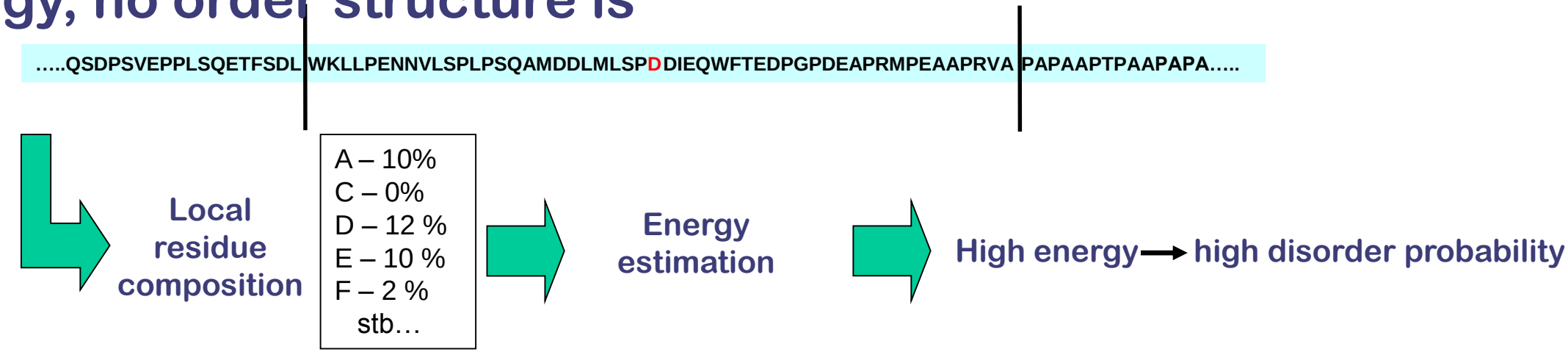
- Physics based approach
  - IUPred
- Machine learning
  - PONDR VSL2
  - DISOPRED3
  - Spritz

  Metaservers

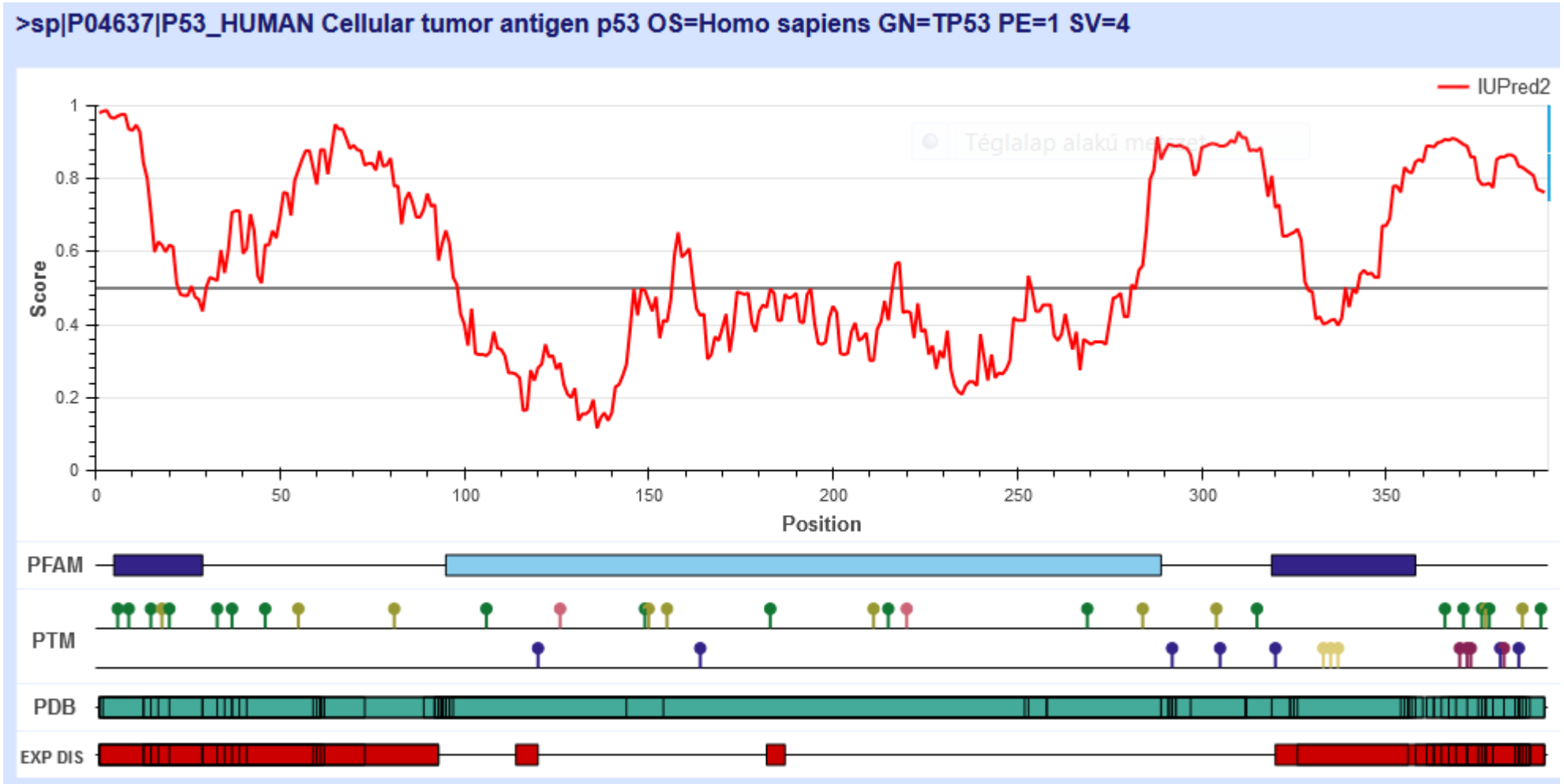
# The IUPred method

- Physics based approach using statistical pairwise potentials based on globular proteins
- transform observed frequencies of amino acid pair interactions into energy-like functions based on the Boltzmann statistics
- Idea: if the neighbour residues do not contribute to a low energy, no order structure is

A – 7.67%
C – 2.43%
D – 4.92 %
E – 5.43 %
F – 3.19 %
…

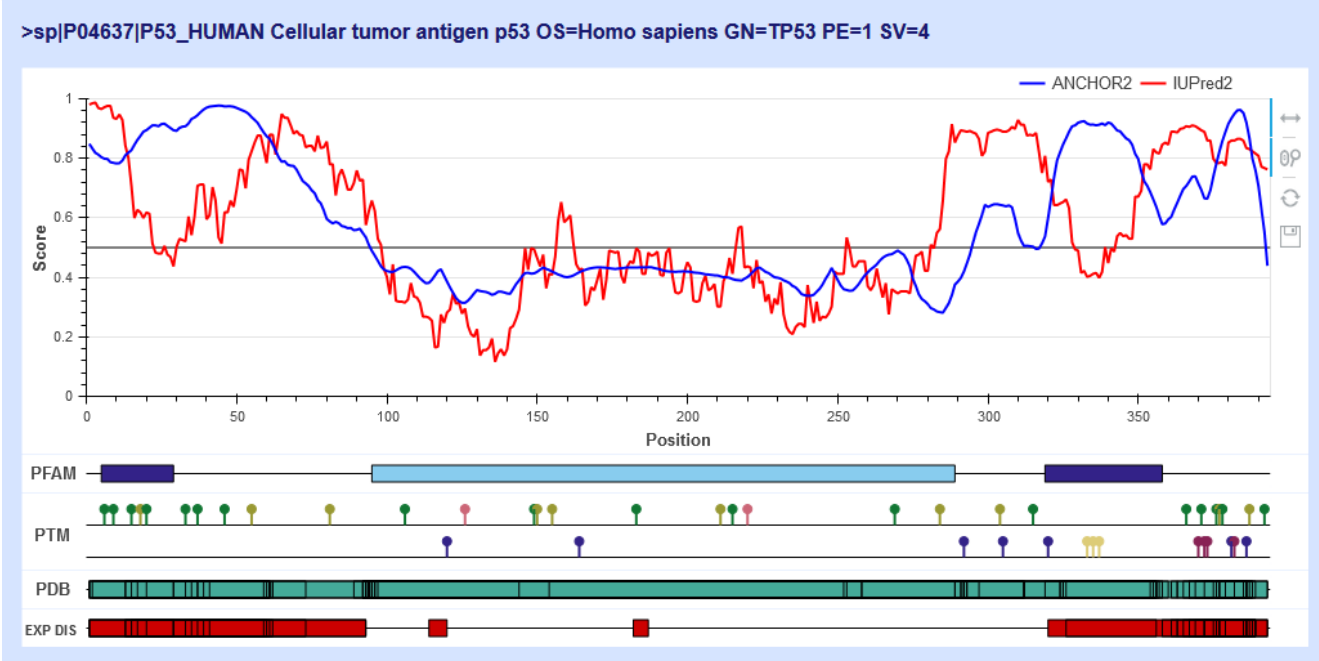


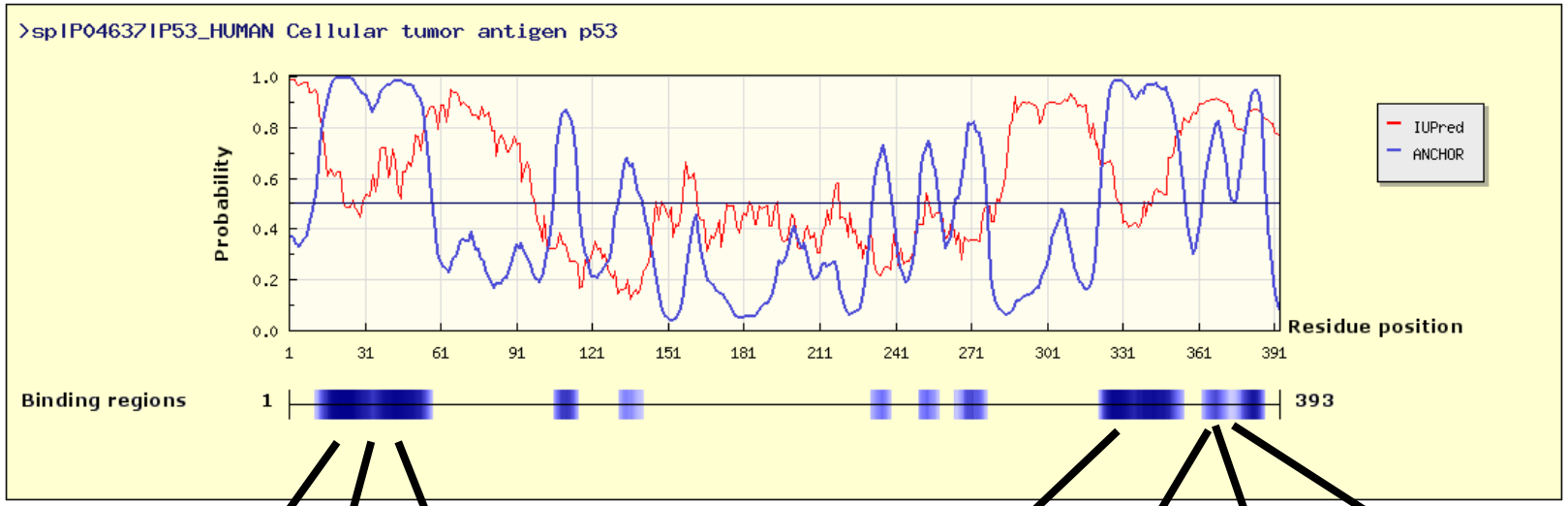
# Output

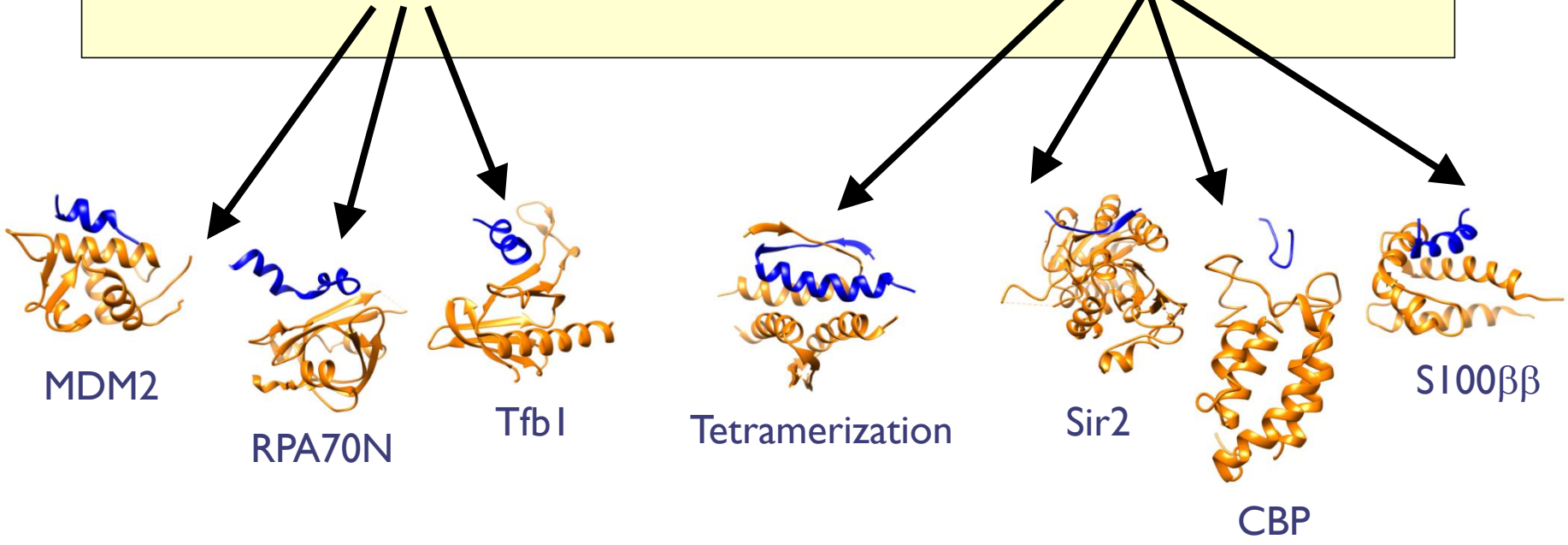

>sp|P04637|P53_HUMAN Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4
IUPred2
Score
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
300
350
Position
PFAM
PTM
PDB
EXP DIS

# Anchor

- Prediction of disordered binding sites in globular proteins
- If the calculated energy if more favourable in a globular environment → binding site




>sp|P04637|P53_HUMAN Cellular tumor antigen p53
Probability
1.0
0.8
0.6
0.4
0.2
0.0
1
31
61
91
121
151
181
211
241
271
301
331
361
391
Residue position
IUPred
ANCHOR
Binding regions
1
393


MDM2
RPA70N
Tfb1
Tetramerization
Sir2
CBP
S100ββ

# Mutual Folding Induced by Binding

Mutual Folding
Induced by Binding

MFIB

Home | Browser | ProteinMap | Search | Downloads | Statistics | Help

- General Information
- Function and Biology
- Structure Summary
  - Chain A
  - Chain B
- Evidence
- Related Structure(s)

**Structural Classification**

**Class:**
Other

**Subclass:**
Ribbon-helix-helix (RHH)

## General Information

**Database accession:** MF2140001

**Name:** Arc repressor

**PDB ID:** 1arq PDB

**Experimental method:** NMR

**Assembly:** homodimer

**Source organism:** *Enterobacteria phage P22*

**Primer publication of the structure:**

***Bonvin AM, Vis H, Breg JN, Burgering MJ, Boelens R, Kaptein R***
*Nuclear magnetic resonance solution structure of the Arc repressor using relaxation matrix calculations.*
(1994) J. Mol. Biol. **236**: 328-41

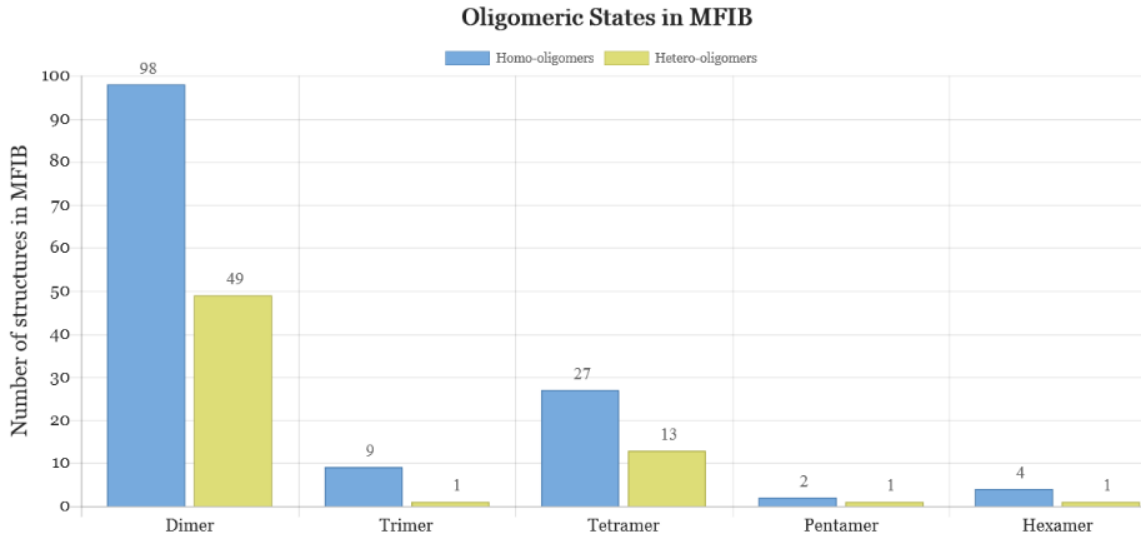
PMID: 8107113 PubMed

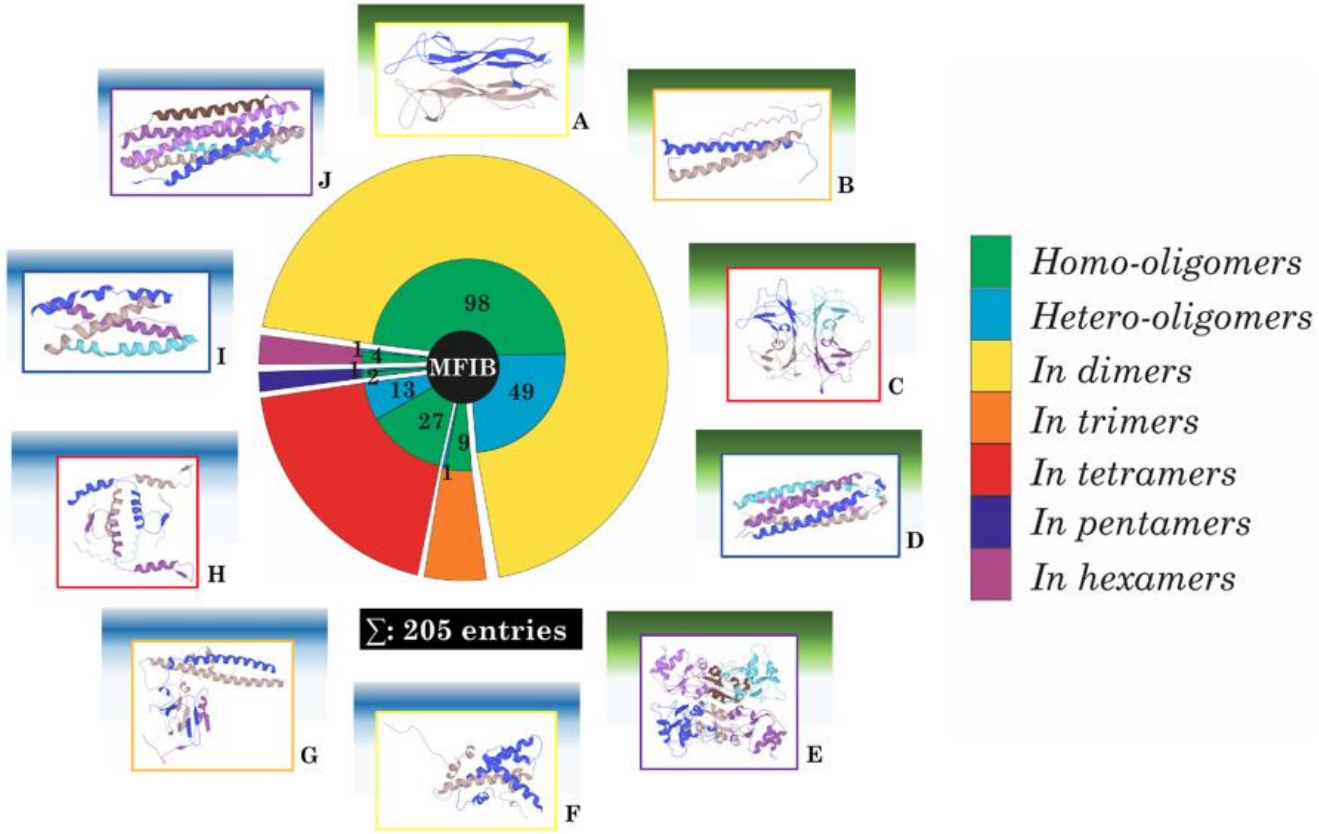
**Abstract:**

The Arc repressor of Salmonella bacteriophage P22 is a dimeric sequence-specific DNA-binding protein. The solution structure of Arc has been determined from 2D NMR data using an "ensemble" iterative relaxation matrix approach (IRMA) followed by direct NOE refinement with DINOSAUR. A set of 51 structures was generated with distance geometry and further refined with a combination of restrained energy minimization and restrained molecular dynamics in a parallel refinement protocol. Distance constraints were obtained from an extensive set of NOE build-ups in H2O and 2H2O via relaxation matrix calculations from the ensemble of structures. Methyl group rotation, aromatic ring flaps and internal mobility effects (via order parameters obtained from a free molecular

*The molecule viewer shows the original PDB stucture.*
*Only the first NMR model was loaded.*

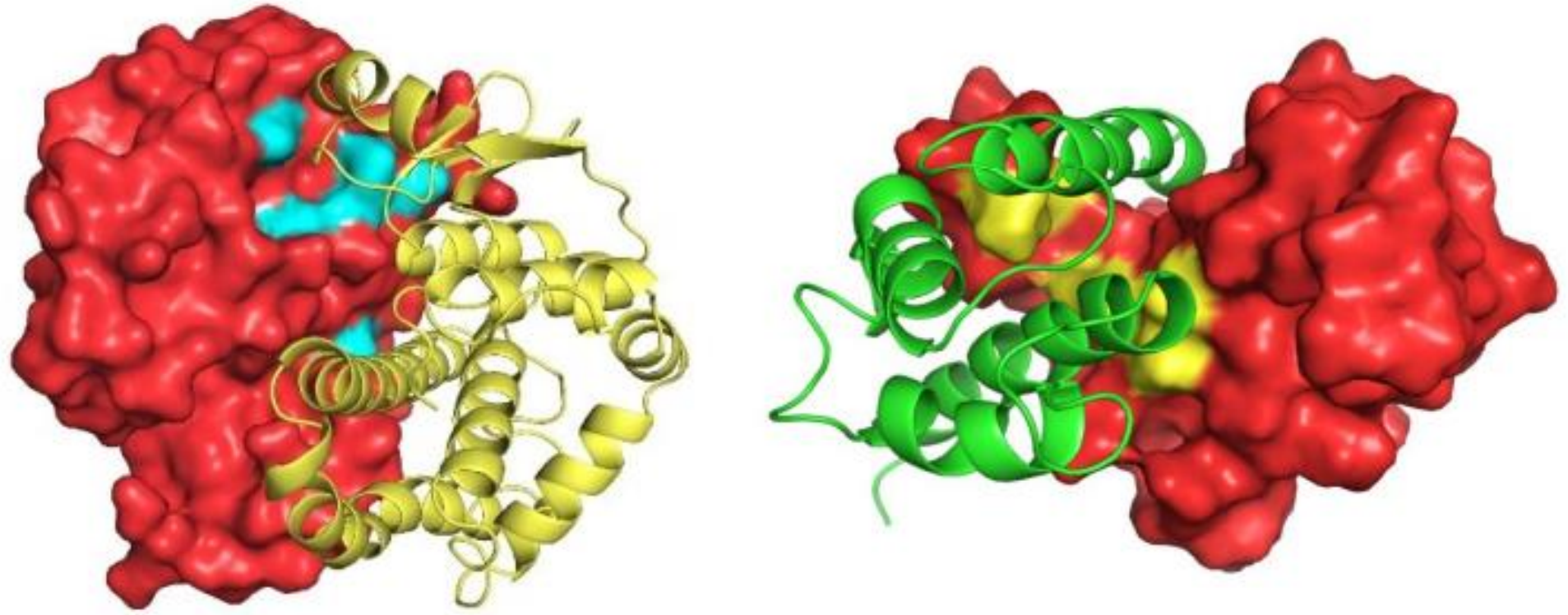
# Oligomeric states in MFIB



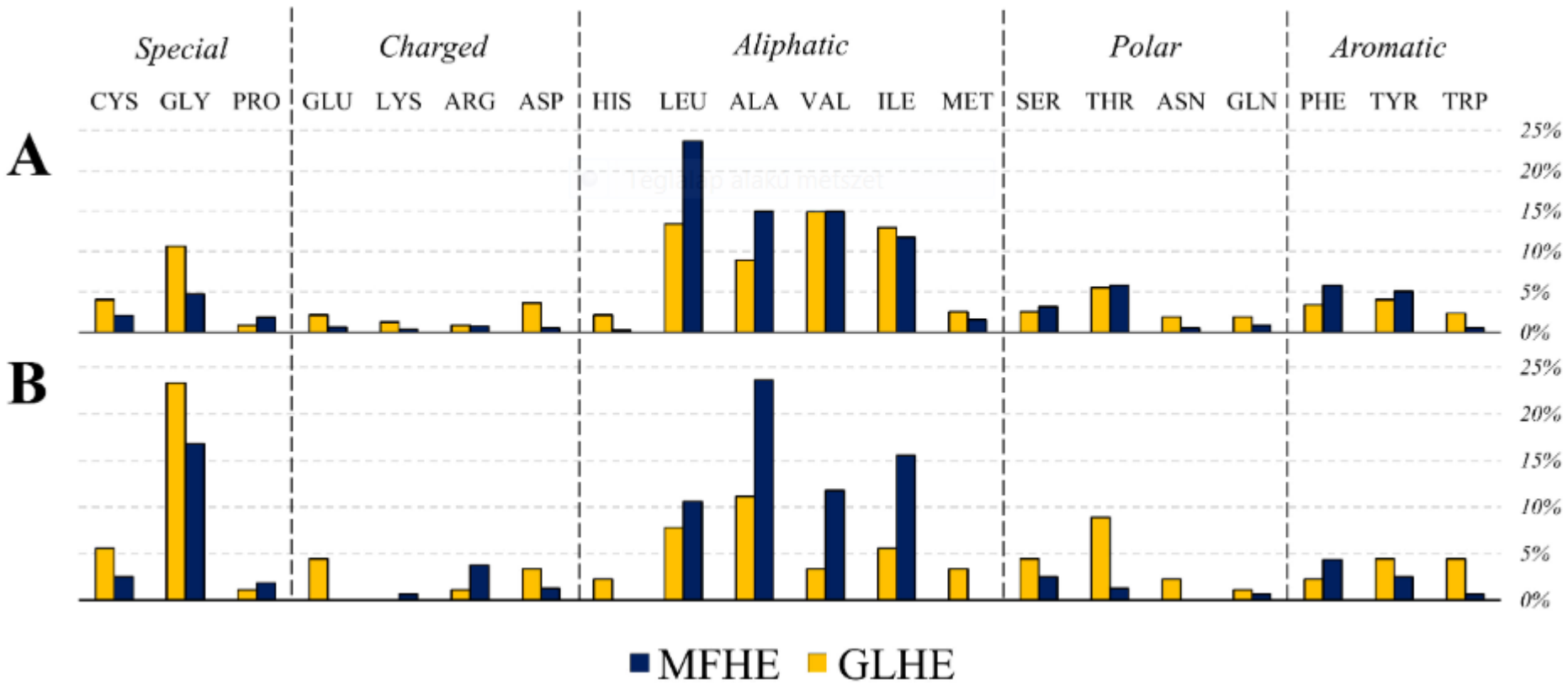



# RSAMP

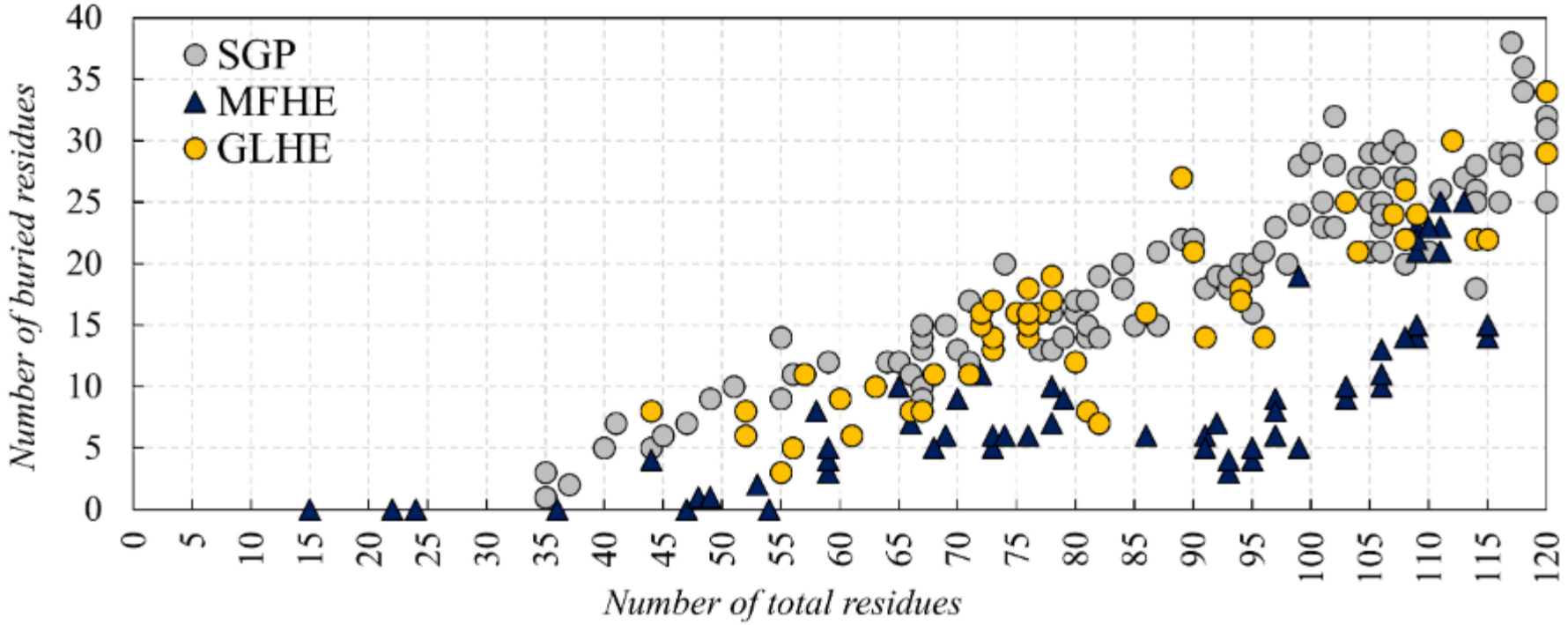
- Residue with Solvent Accessible Main-chain Patches



# Amino acid compositions



# Number of buried residues



# Burial of residues in monomers / dimers