

Elméleti módszerek lipidek és membránfehérjék tanulmányozására

Hegedűs Tamás

tamas@hegelab.org

MTA-SE Molekuláris Biofizikai Kutatócsoport
SE Biofizikai és Sugárbiológiai Intézet



Fehérjék szerkezetének és dinamikájának jelentősége

A betegség molekuláris szintű oka?
A gyógyszer-kötő zseb alakja?

37°C-on, oldatban nem egy szerkezet létezik,
hanem egy konformációs sokaság.

Számítógépes modellezés jelentősége

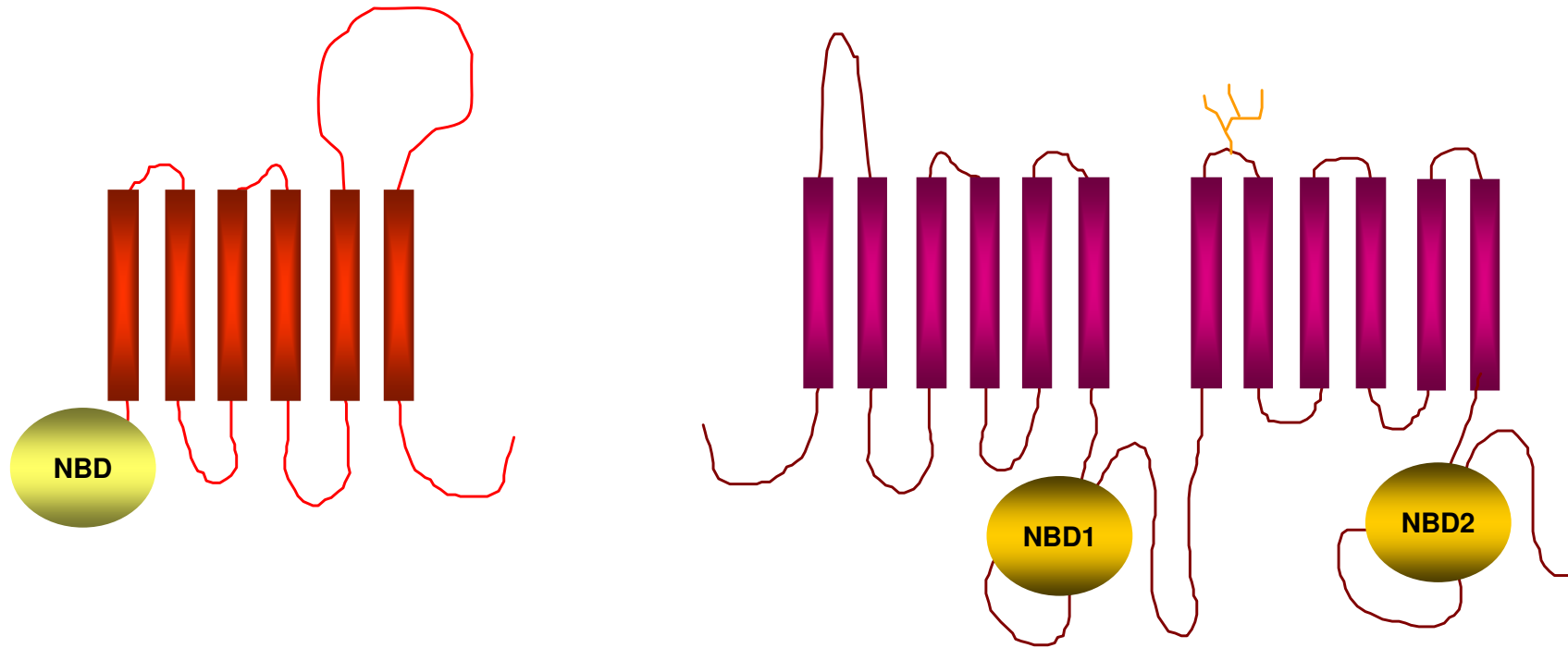
Atomi szintű információt ad mozgásokról.

**Kísérletes módszerek
általában nem szolgáltatnak
közvetlen információt az atomi szintű történésekről.
(pl. EPR, X-links: nem, NMR: igen, de korlátoltan)**

Membránfehérjék szerkezeti modellezése

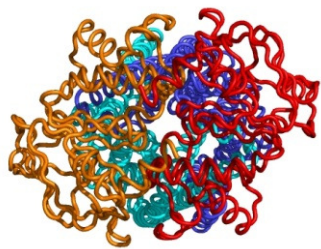
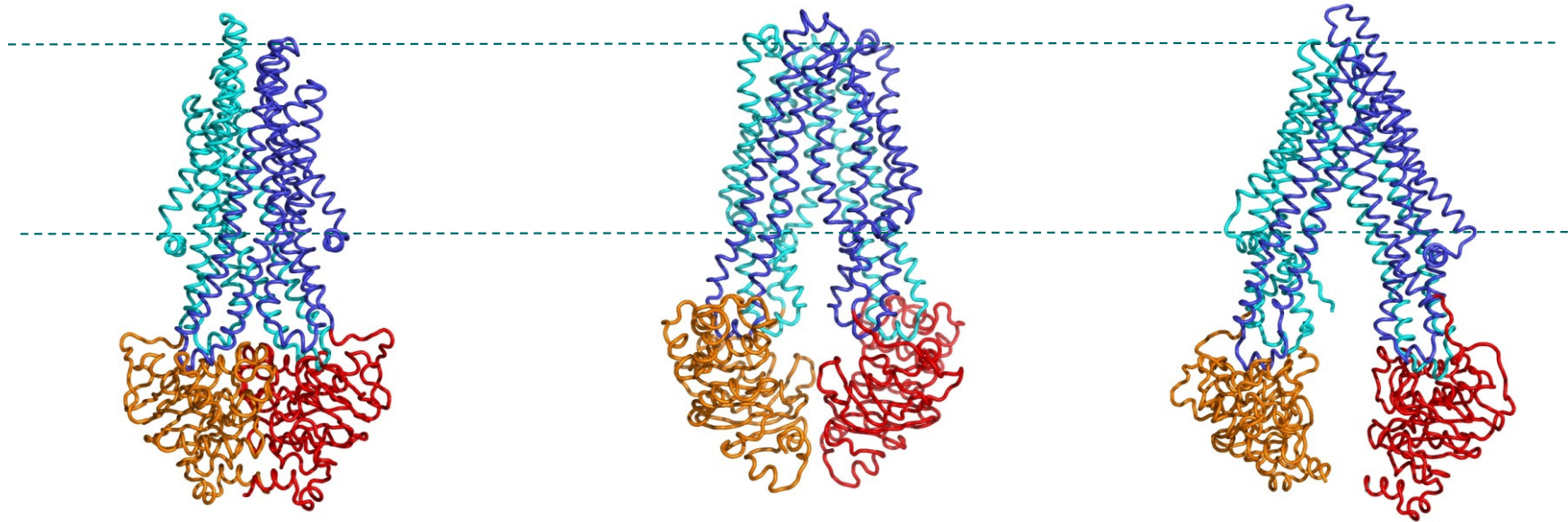
- Kísérleti adatok alapján
- Elméleti módszerekkel
 - homológia modellezés
 - mélytanulásos módszerek

ATP Binding Cassette (ABC) proteins

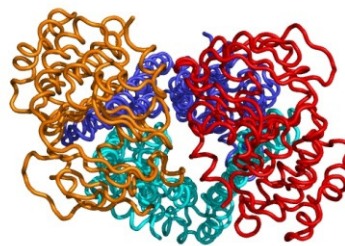


Conformation of ABC proteins

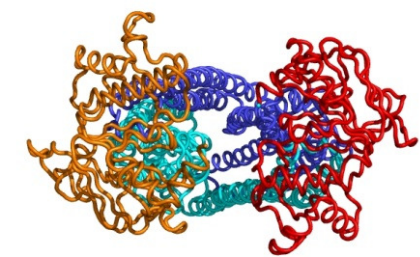
(Pgp-like)



**“bottom-closed”
(+ATP)**

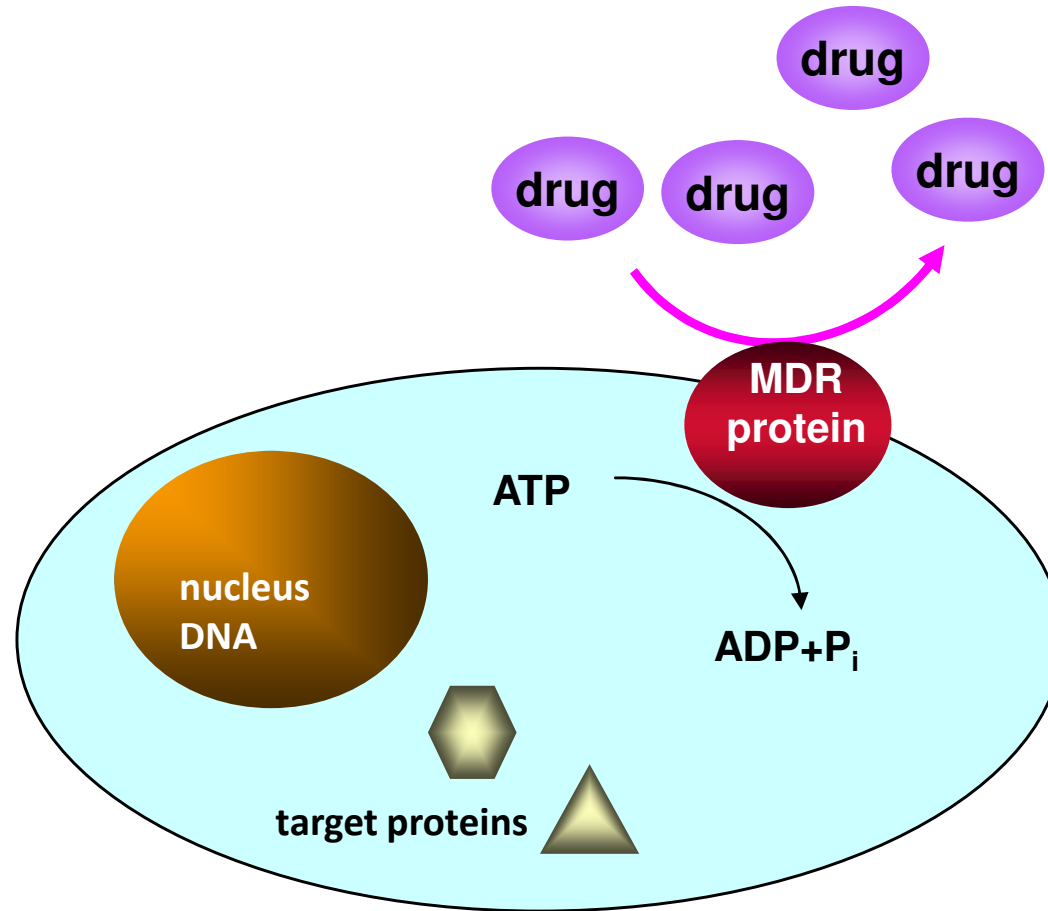


**“bottom-closed” apo
(-ATP)**

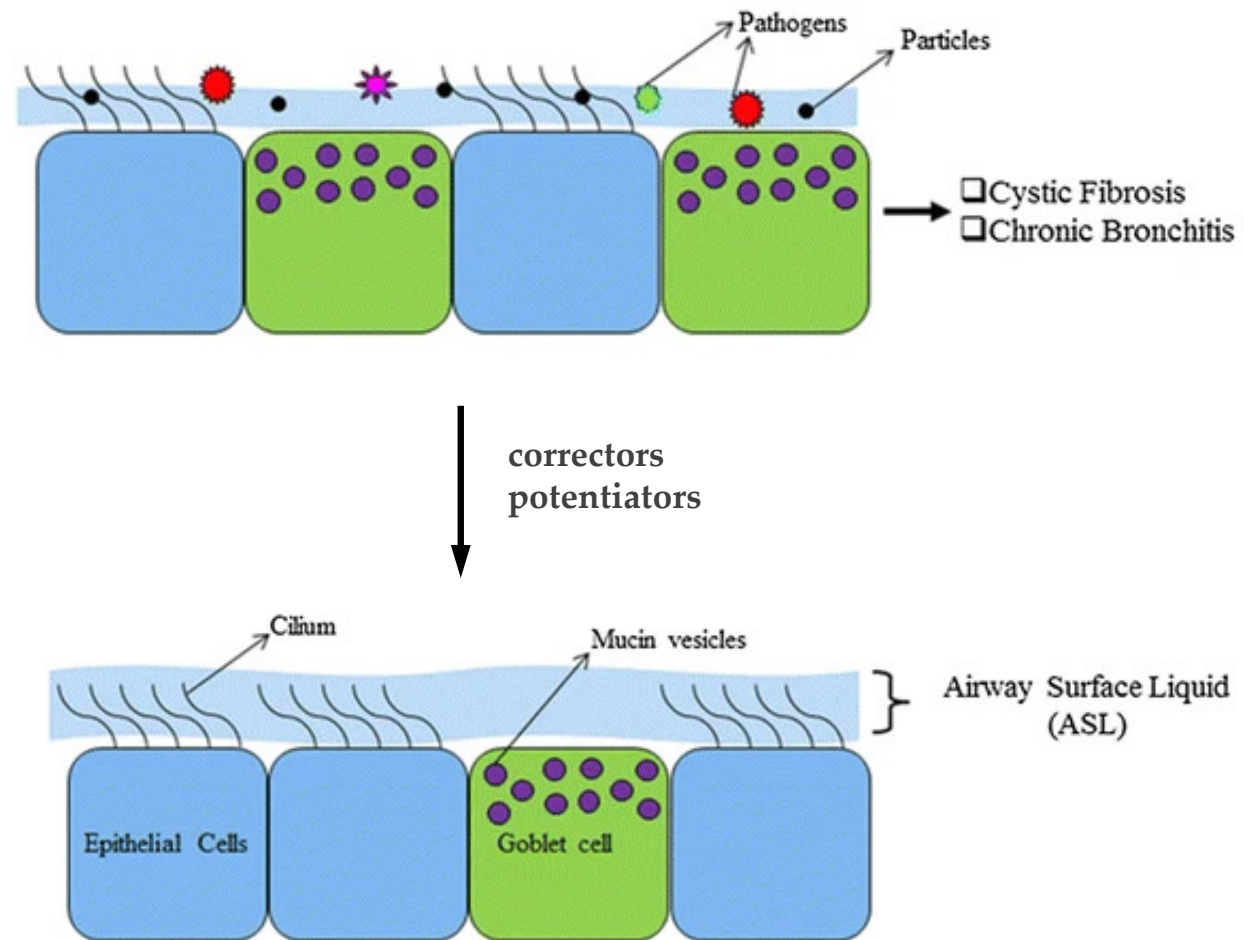


**“bottom-open” apo
(-ATP)**

Multidrug transport – ABC proteins

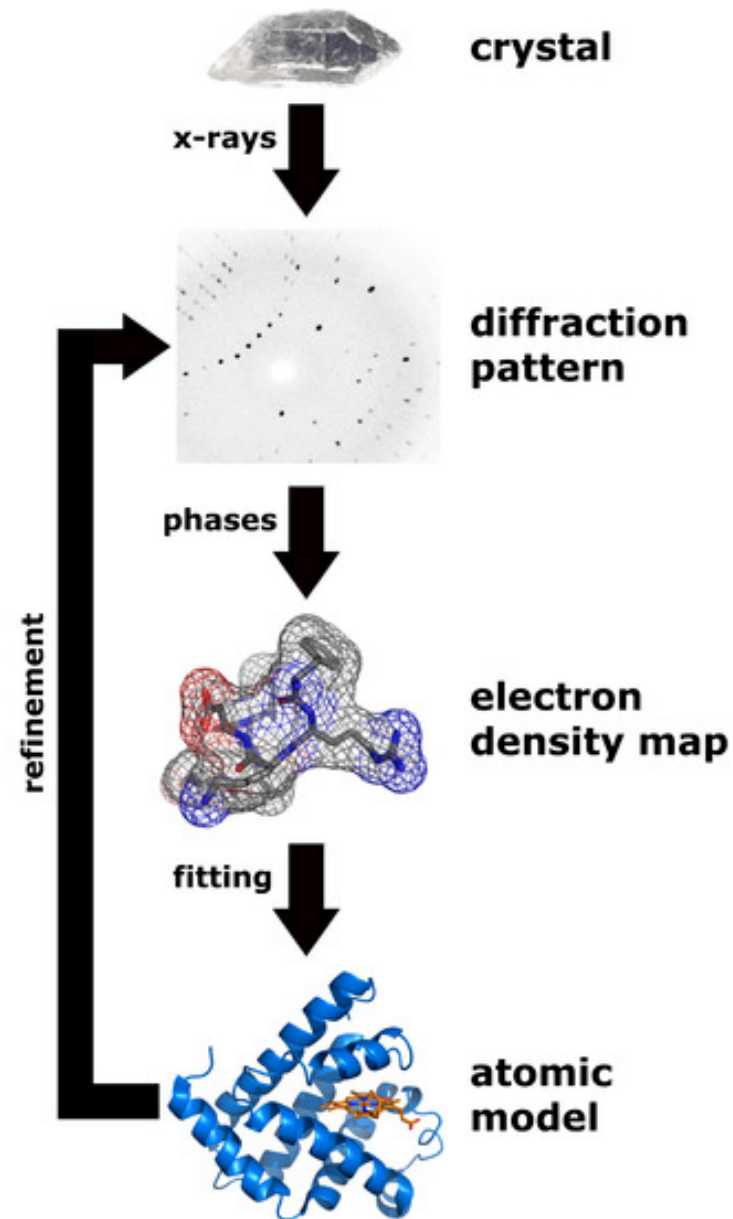


Cystic fibrosis (CF)



Ghosh, Boucher, Tarran,
CMLS 2015

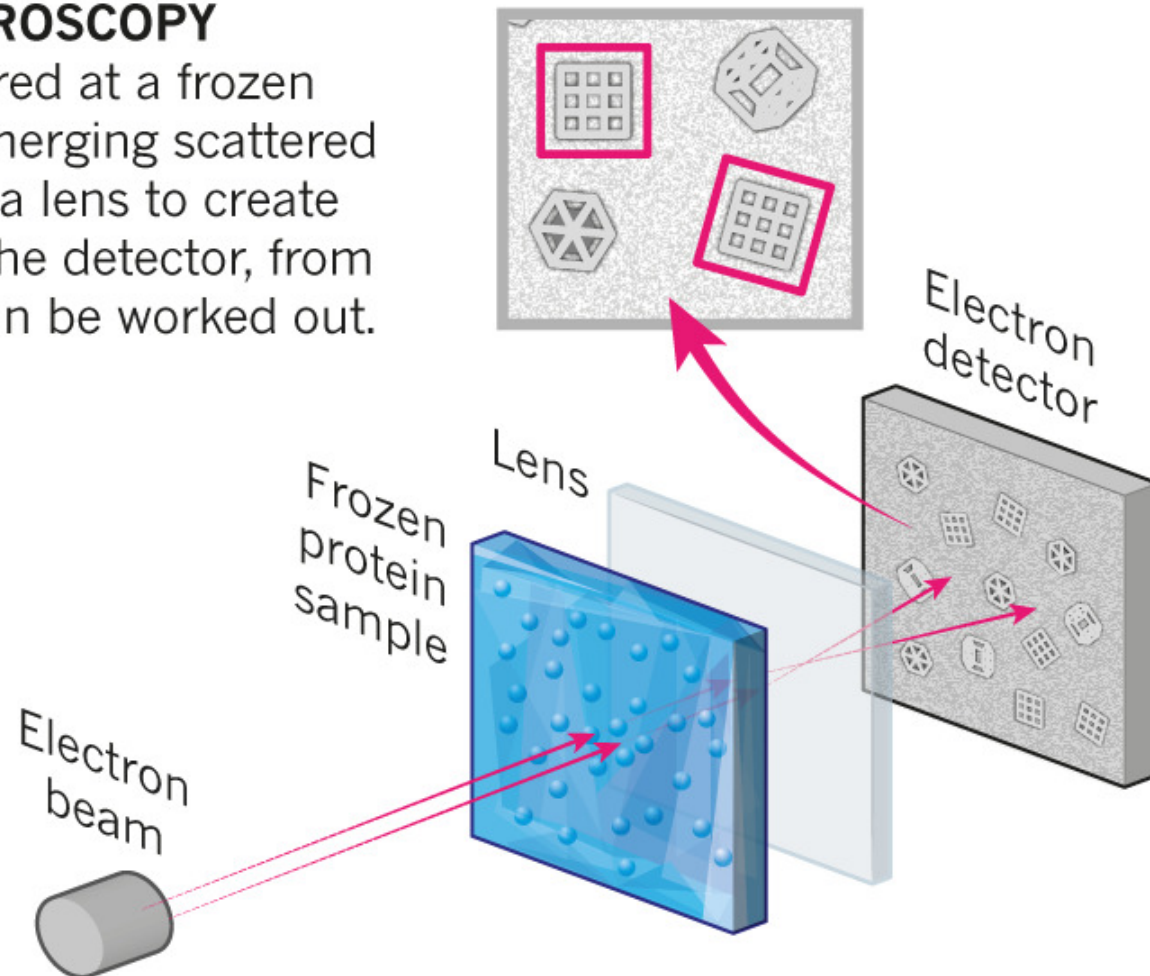
Szerkezet meghatározás – Röntgen kristallográfia



Szerkezet meghatározás – krio-EM

CRYO-ELECTRON MICROSCOPY

A beam of electron is fired at a frozen protein solution. The emerging scattered electrons pass through a lens to create a magnified image on the detector, from which their structure can be worked out.



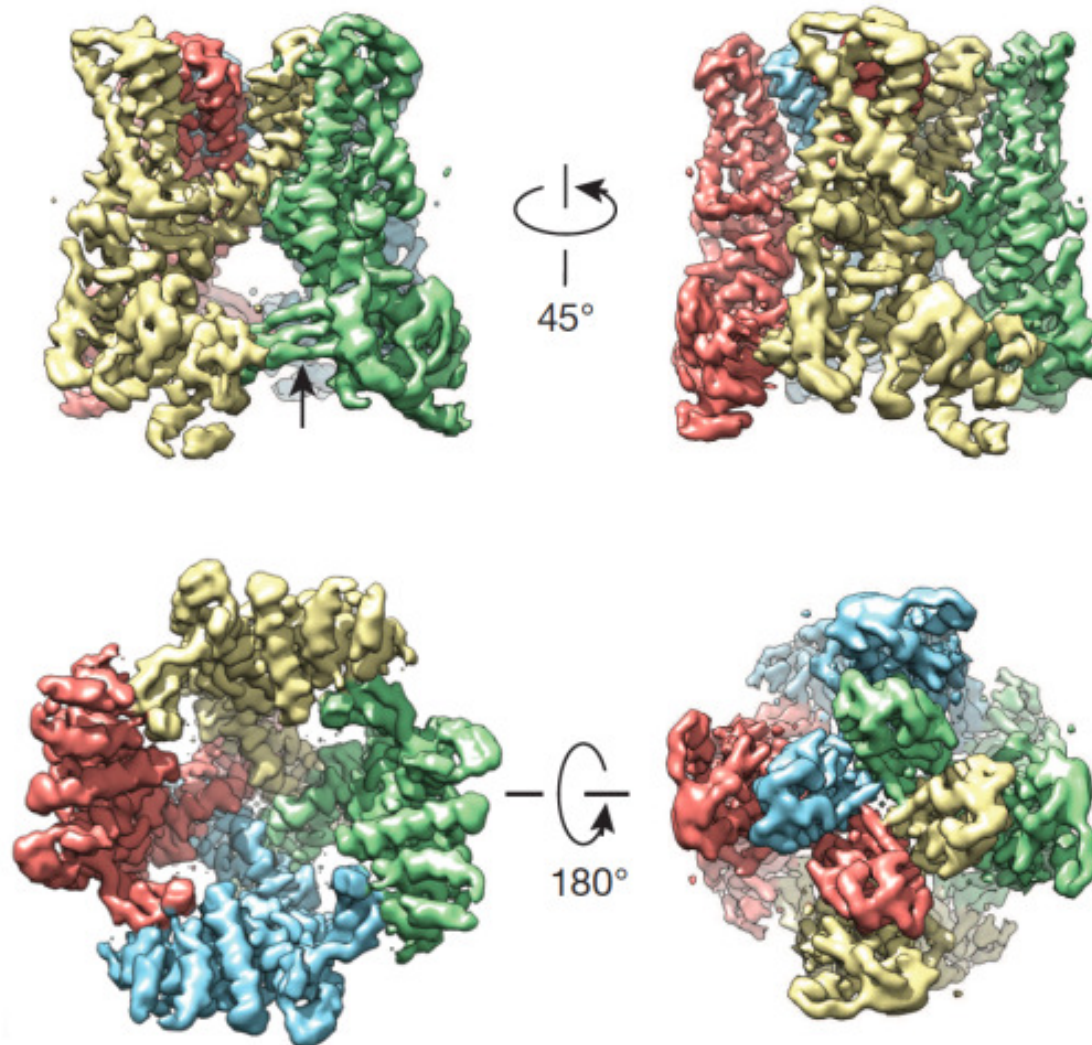
© nature

Ewen Callaway, Nature | News Feature

The revolution will not be crystallized: a new method sweeps through structural biology, 09 September 2015

Szerkezet meghatározás – „single particle”

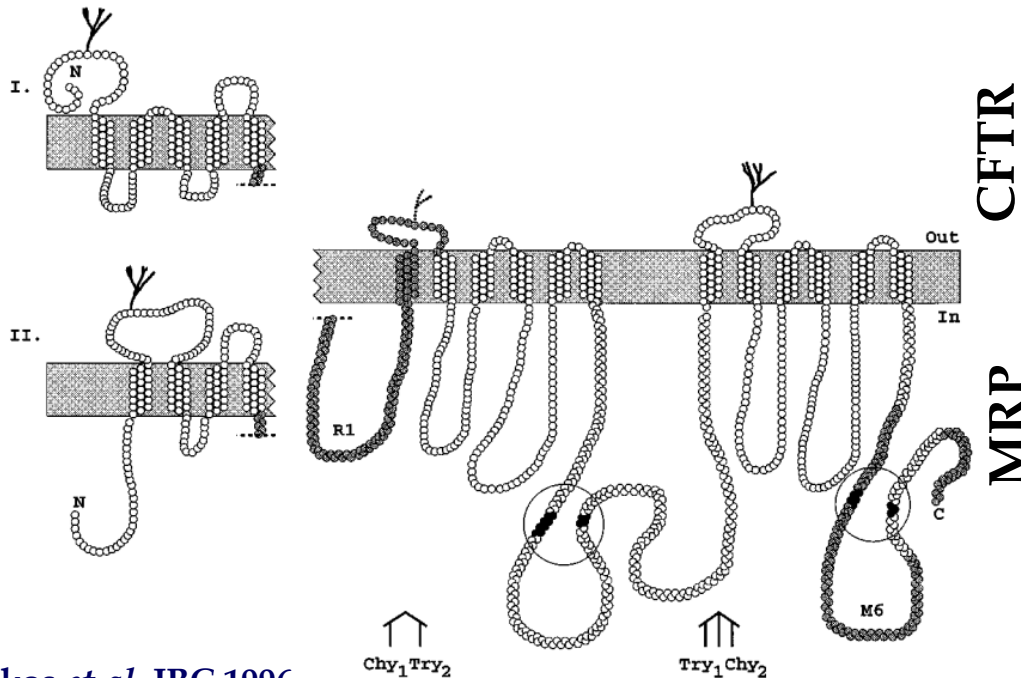
The TRPV1 channel detects the burn of chilli peppers, and this 3.4-Å structure is considered super-hot in the structural-biology world.



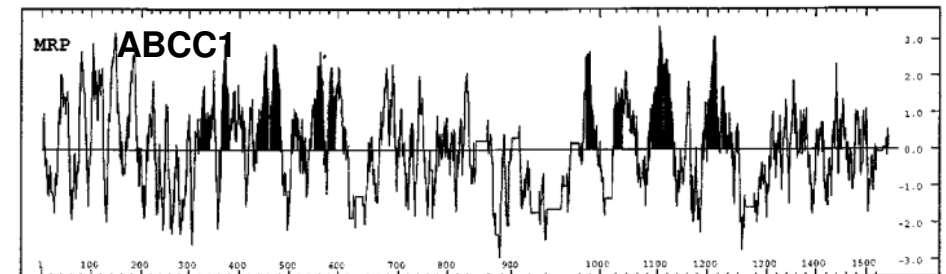
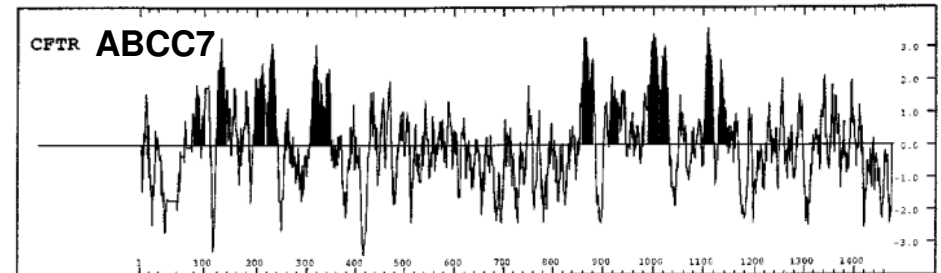
Membránfehérjék topológiája

MRP

Hidrofóbicitási görbék:



Bakos *et al.* JBC 1996



CFTR topológiája kísérletekből ismert

Chang *et al.* J Biol Chem. 1994 269(28):18572-5

Kísérletes topológia meghatározási módszerek:

- Glikozilációs hely és mintázat
- *tag* – felismerés
- Cys hozzáférhetőség

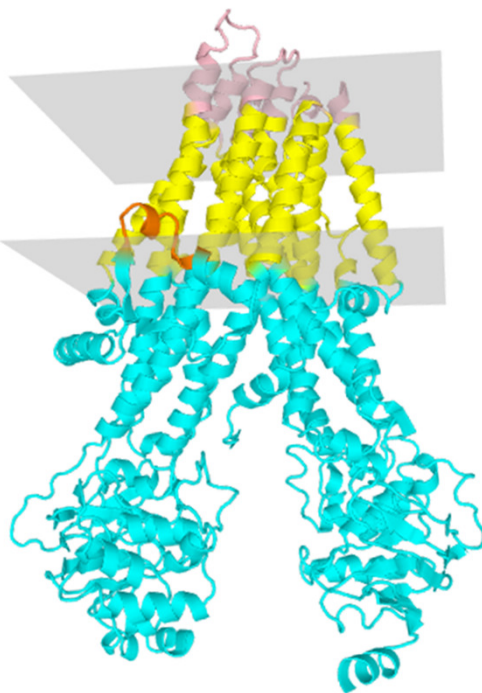
Membrántopológia és TM hélixek predikciója szekvencia alapján

- Szekvencia alapján
 - *positive-inside rule*
 - TM hélixek és nem-TM részek aminosav eloszlása más
- Biológiai tudás figyelembevételének lehetősége
- Több prediktor integrálása
pl. CCTOP.enzim.hu

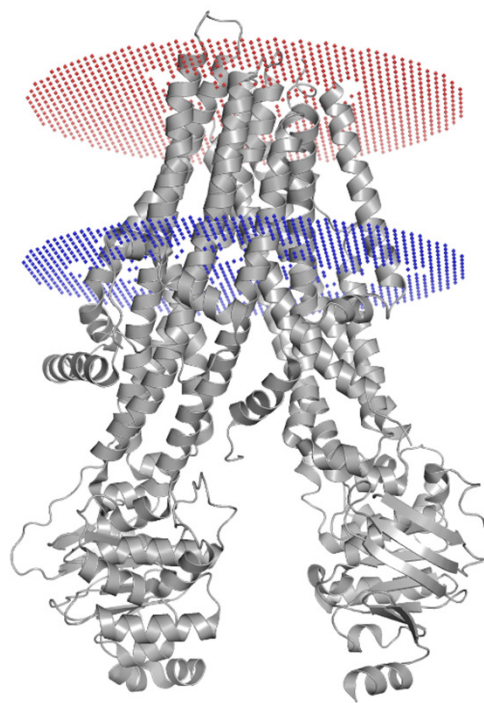
TM hélixek predikciója szerkezet alapján

1. „water accessible area”, lipid kettősréteg lehetséges helye, pontozás – fitness függvény
2. molekula dinamika szimulációk

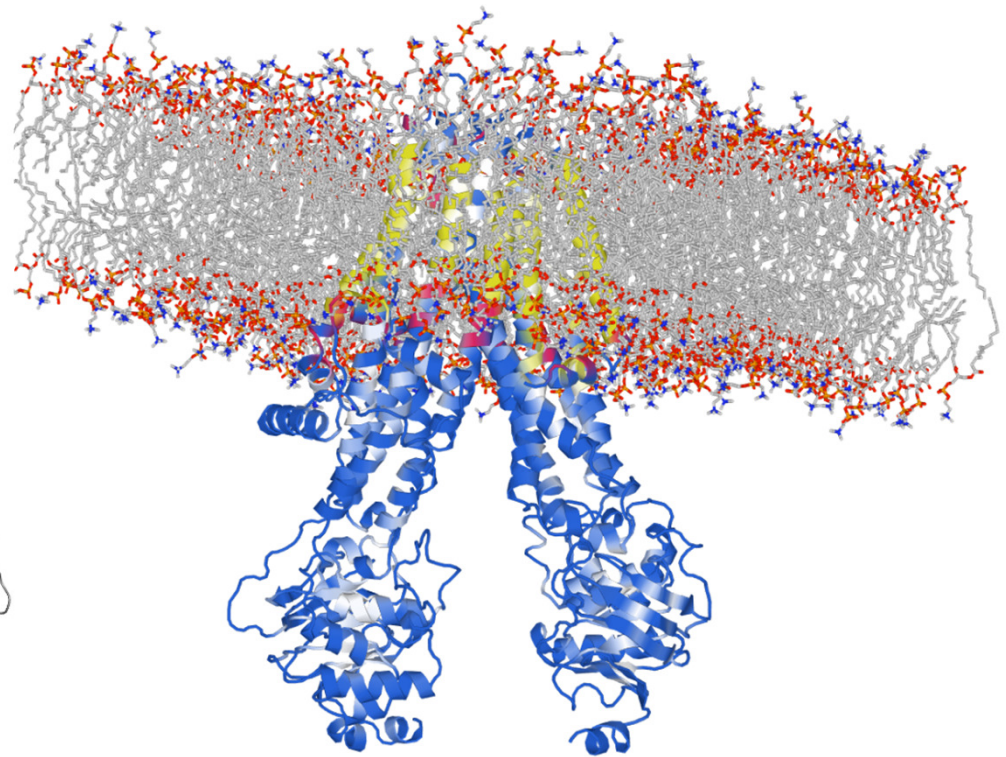
TMDet
PDBTM



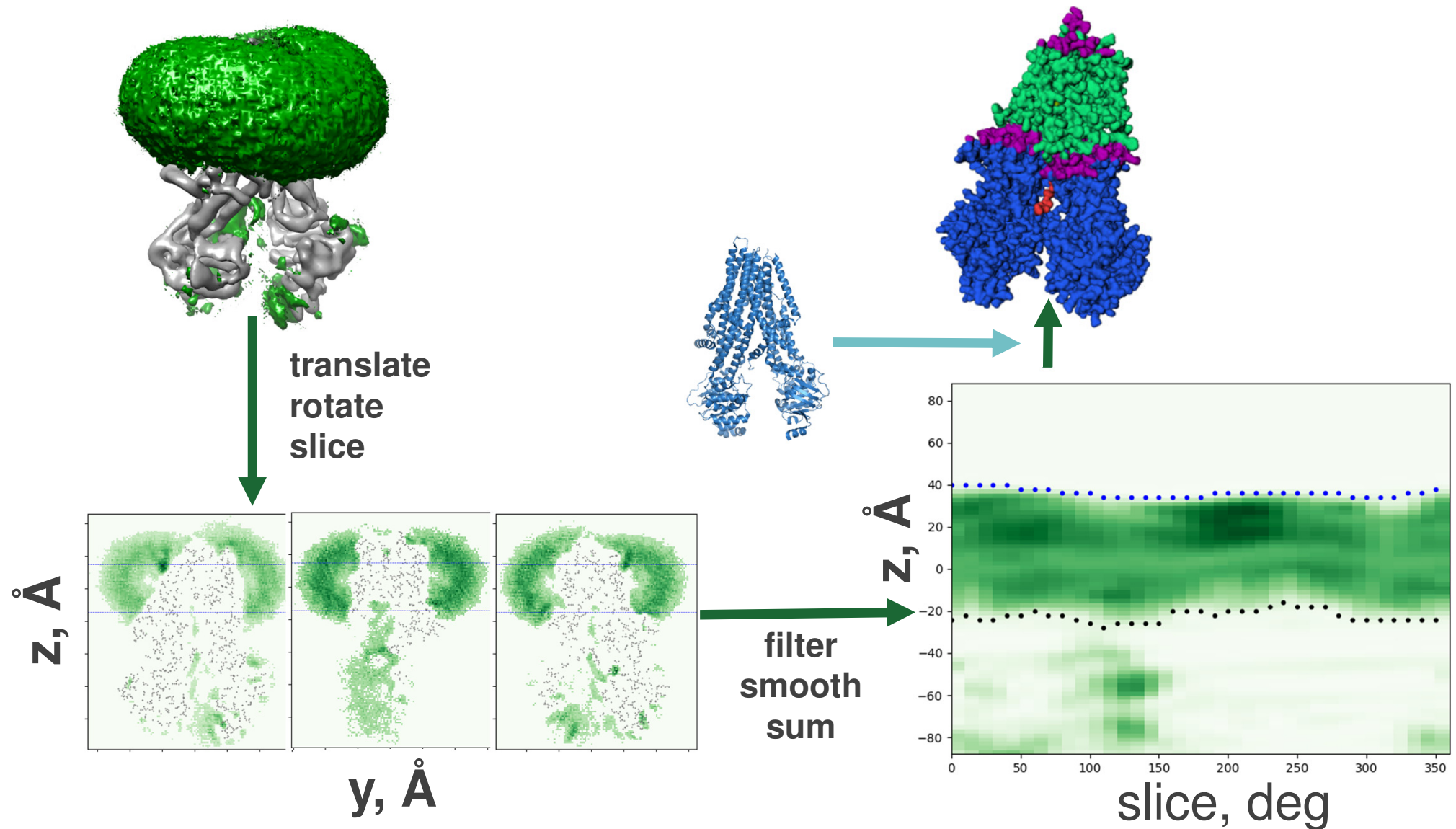
PPM
OPM



MemProtMD



The MemBlob can be converted to membrane boundaries



Homológia modellezés I.

- **Templát keresés**
- **Szekvencia illesztés**
- **Modellezés**
- **Energia-minimalizálás**

BLOSUM (BLOCKs of Amino Acid SUBstitution Matrix) matrix is a substitution matrix

BLOSUM (BLOCKs of Amino Acid SUBstitution)
is a substitution matrix

Ala	4																					
Arg	-1	5																				
Asn	-2	0	6																			
Asp	-2	-2	1	6																		
Cys	0	-3	-3	-3	9																	
Gln	-1	1	0	0	-3	5																
Glu	-1	0	0	2	-4	2	5															
Gly	0	-2	0	-1	-3	-2	-2	6														
His	-2	0	1	-1	-3	0	0	-2	8													
Ile	-1	-3	-3	-3	-1	-3	-3	-4	-3	4												
Leu	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4											
Lys	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5										
Met	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5									
Phe	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6								
Pro	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7							
Ser	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4						
Thr	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5					
Trp	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11				
Tyr	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7			
Val	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4		
Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val			

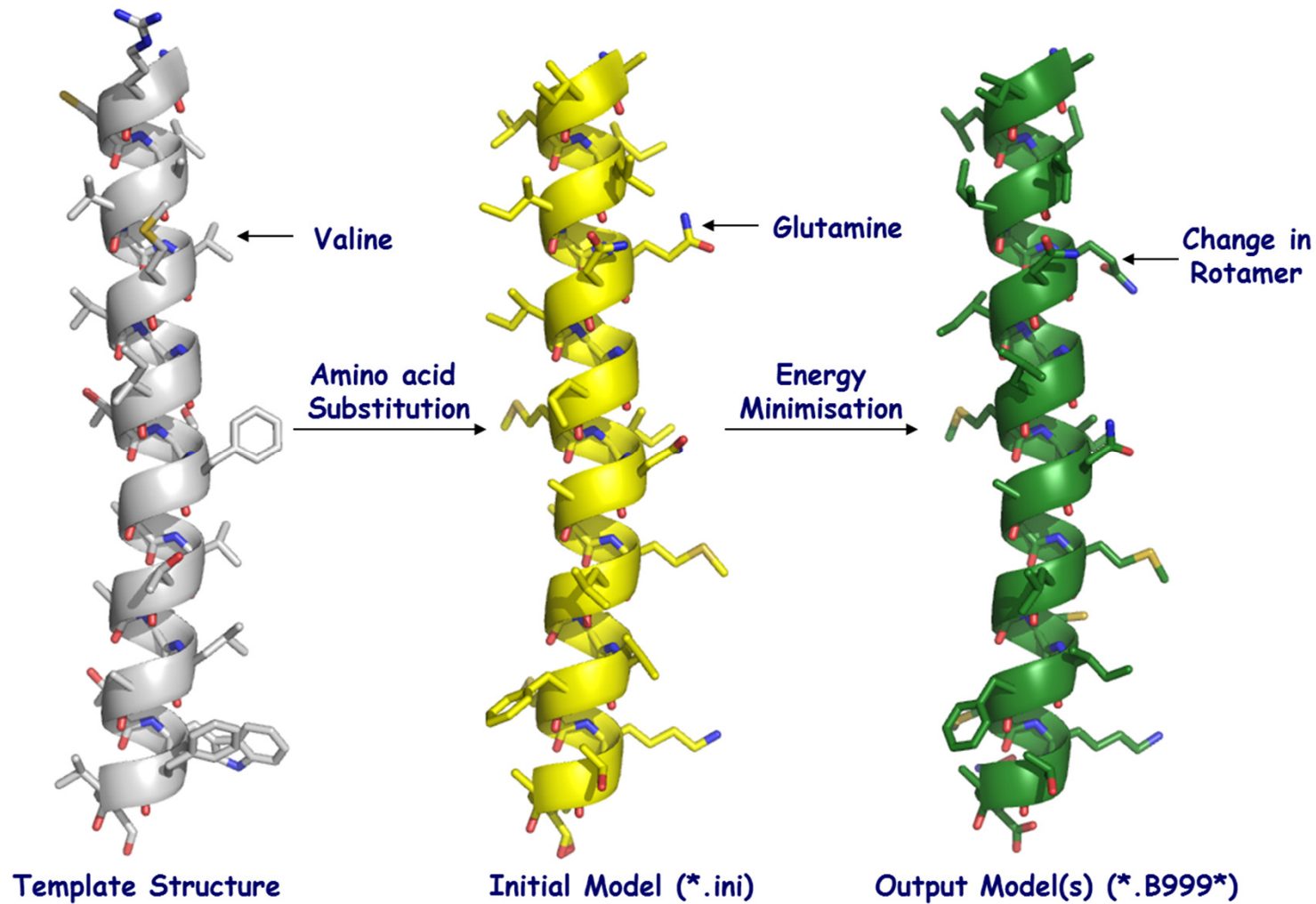
CLUSTAL W (1.83) multiple sequence alignment

Alignement – pl. ClustalW

2HYD -----MIKRYLQFVK-----PYKYRIFATIIVGIIKFGLPMLIP
3B5X -----WQTFKRLWTYIR-----LYKAGLVVSTIALVINAAADTYMI
CFTR_HUMAN MQRSPLKASVVSKLFFSWTRPIILRKGYRQRLELSDIYQIPSVDSDNLS

* : * : :

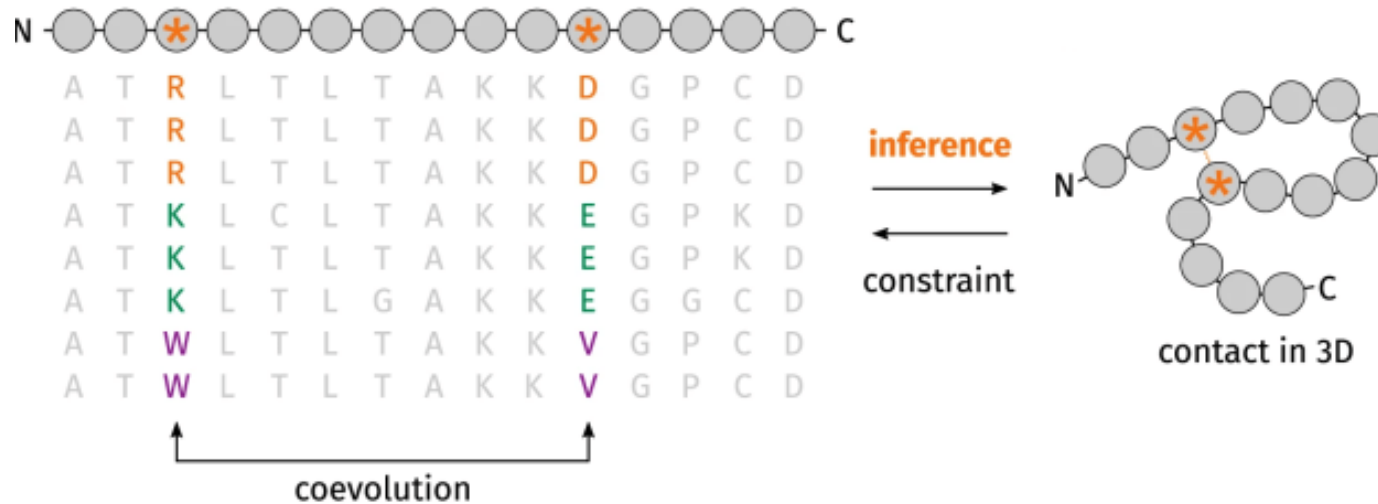
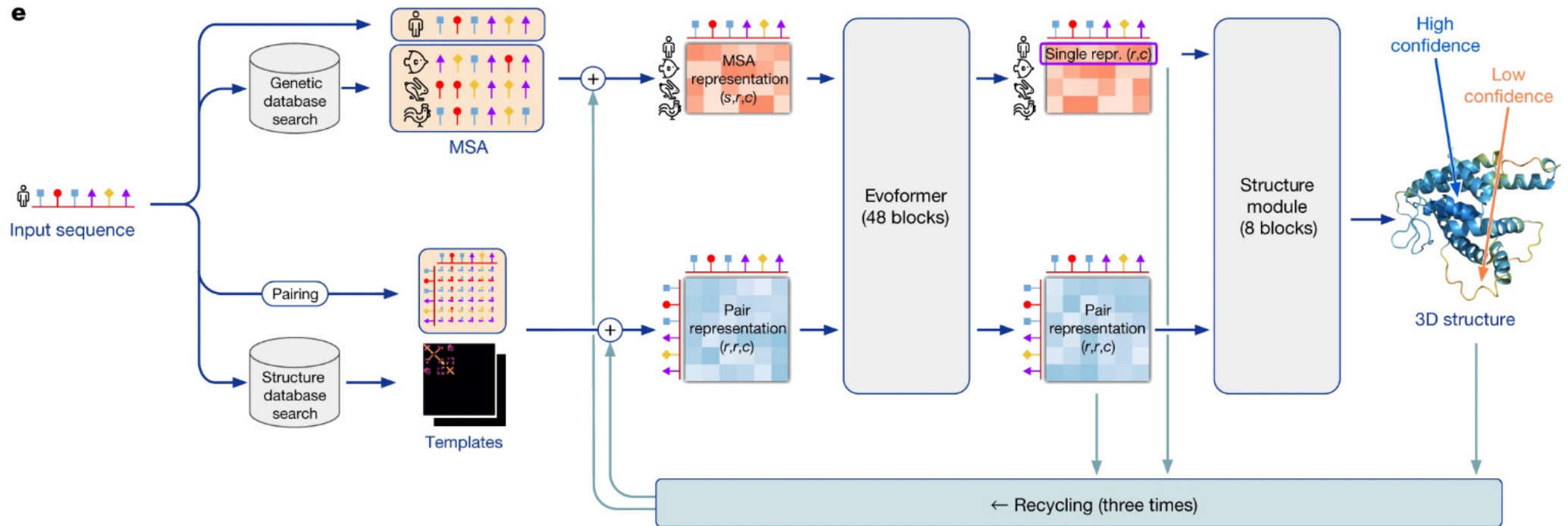
Homológia modellezés II.



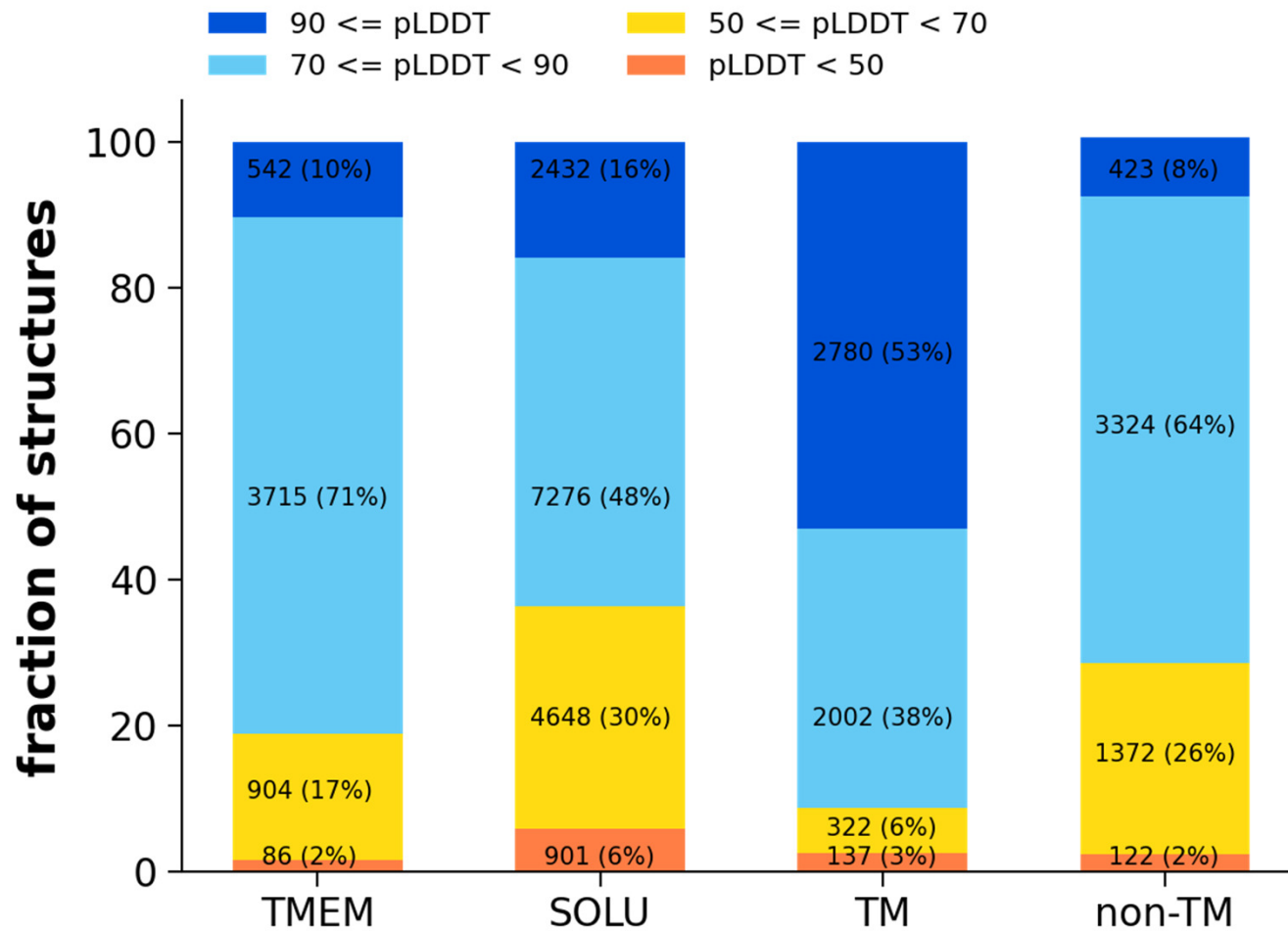
AlphaFold2

machine learning, deep learning, AI

Jumper et al. Nat 2021

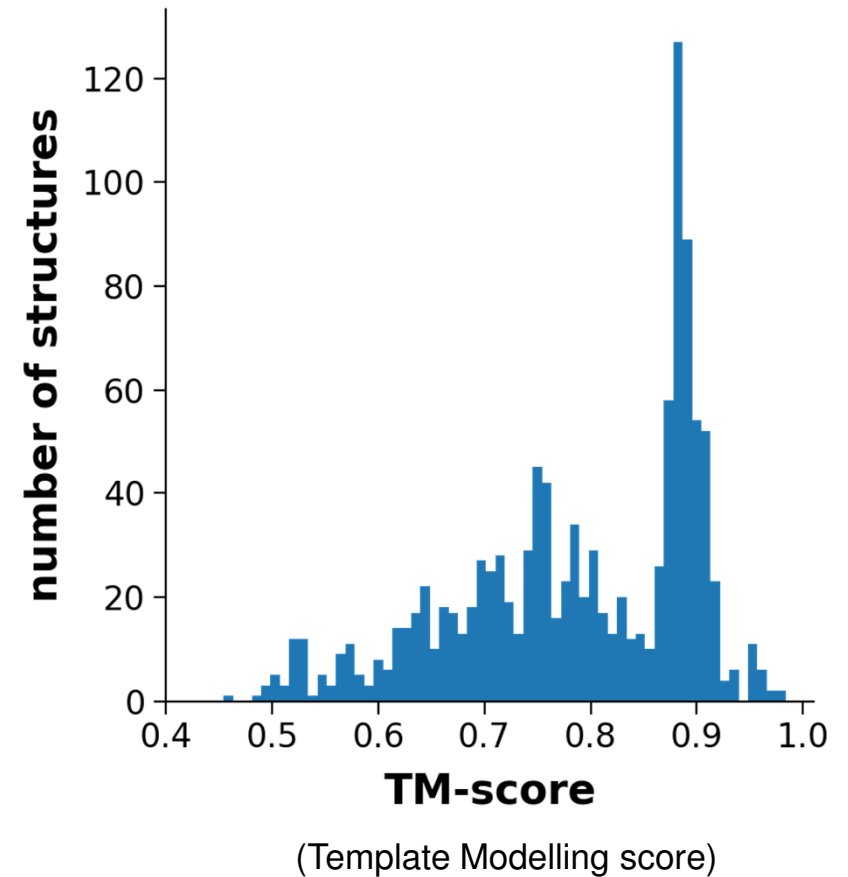
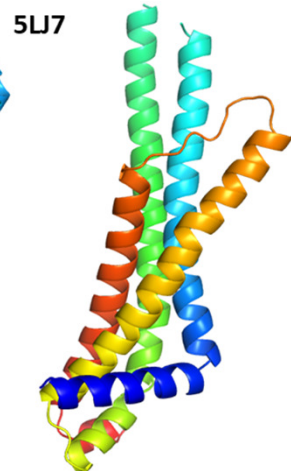
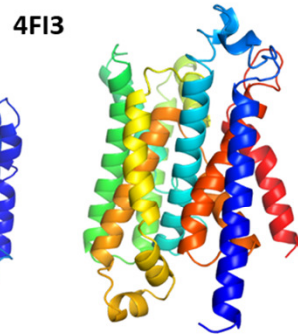
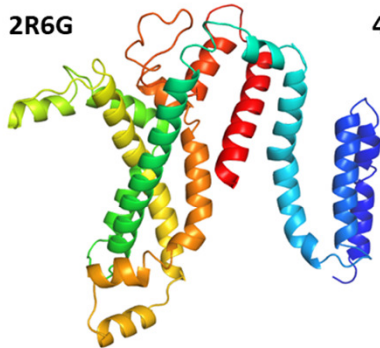
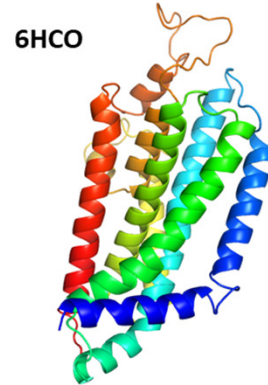


TM protein structure prediction by AF2



ABC protein folds

fold class	reference PDB
Pgp-like	4M1M
ABCG2-like	6HCO
MalFG-like	2R6G
BtuC-like	4FI3
EcfT-like	4HUQ
LptFG-like	5X5Y
MacB-like	5LJ7
MlaE-like	7CH0



CFTR TM8

Farkas *et al.* CMLS 2019

Hegedus *et al.* CMLS 2022

cryo-EM

1/22 simulations, $r \geq 1.8$

Å

0.36%

cryo-EM

AF2

homo

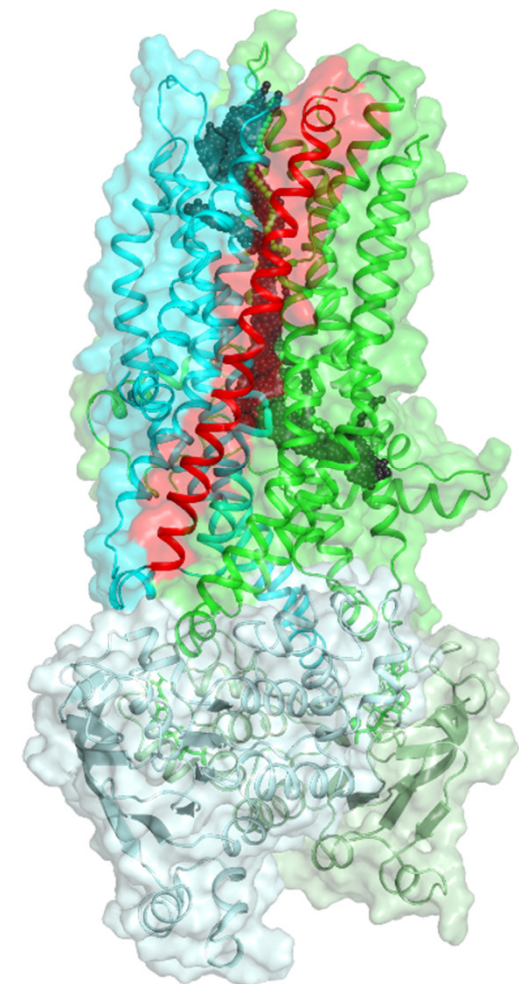
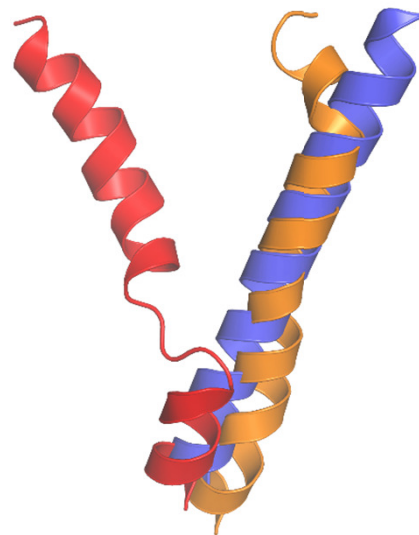
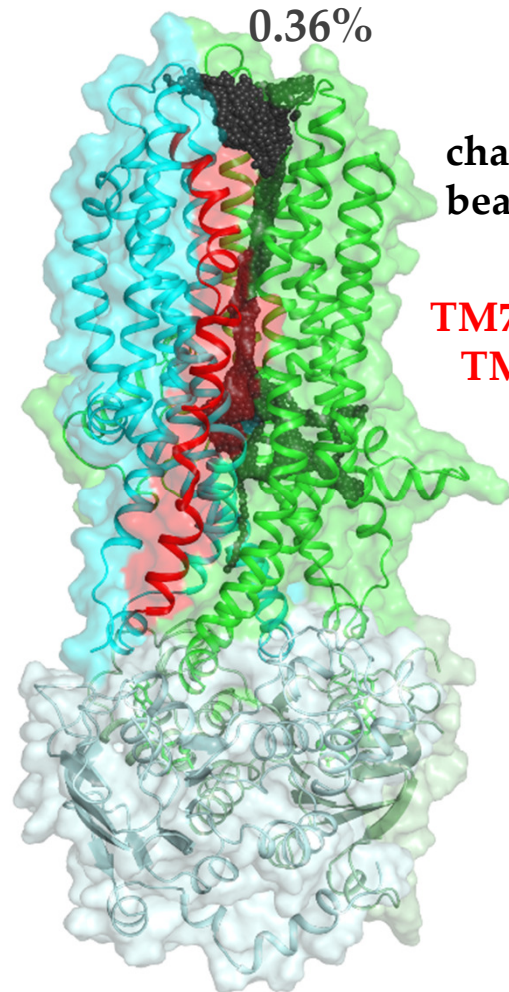
channel
beads

TM7
TM8

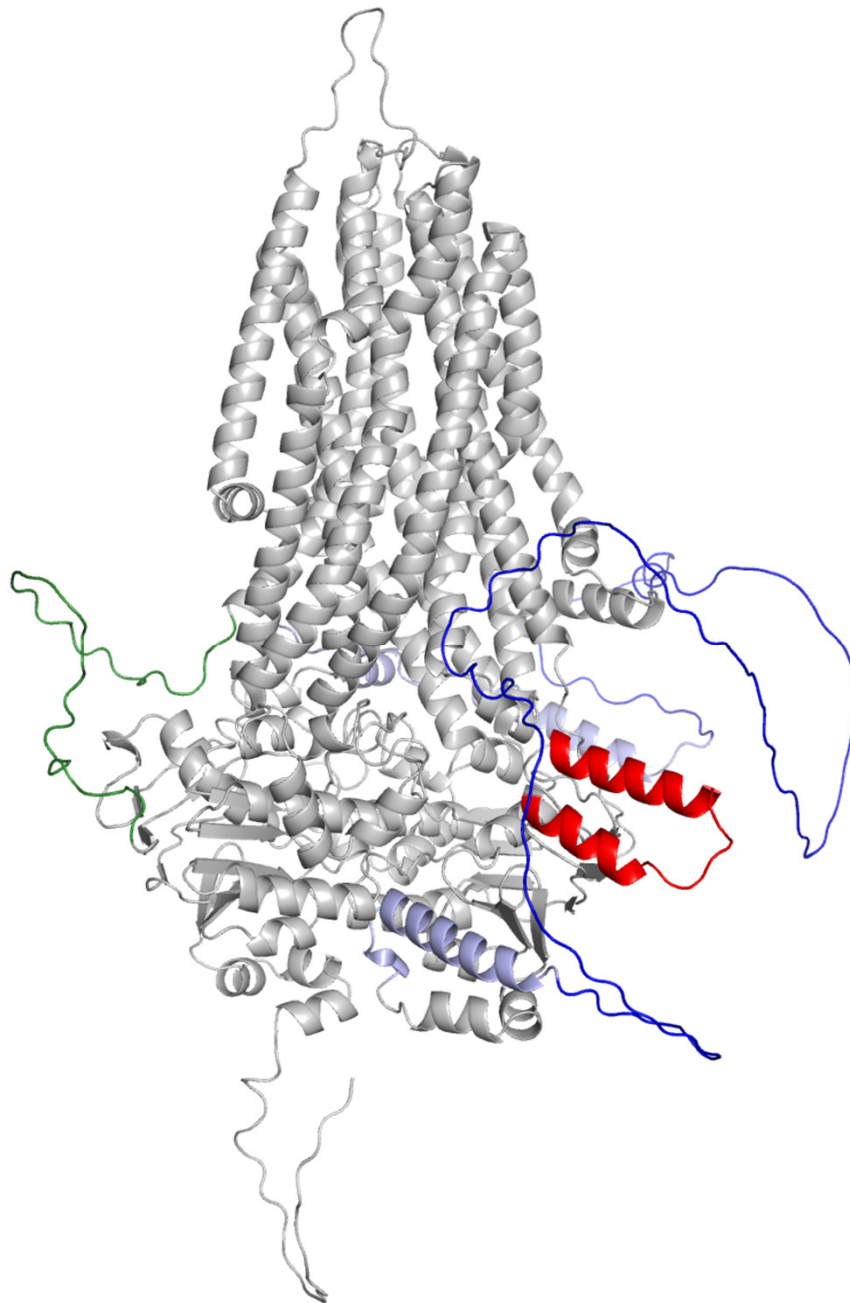
homology modelling

5/6 simulations

3.74%

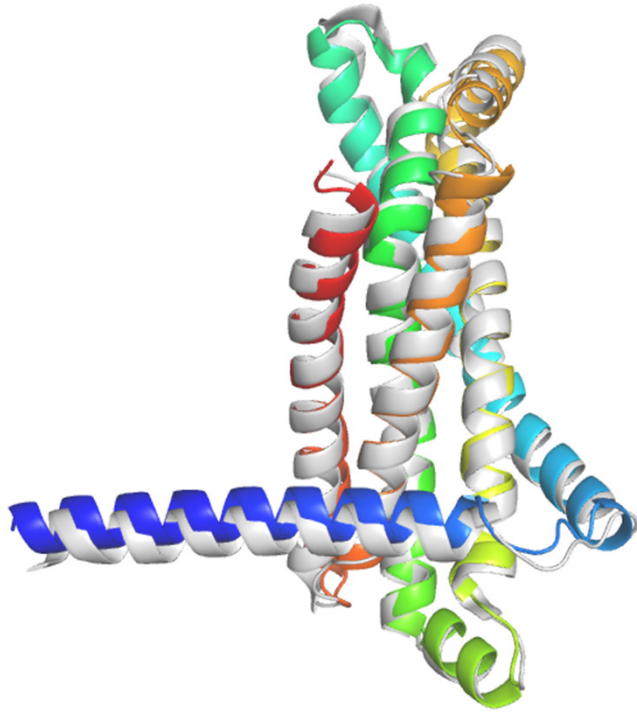


CFTR – disordered regions



pLDDT score - IDR prediction

Prediction of new TM folds



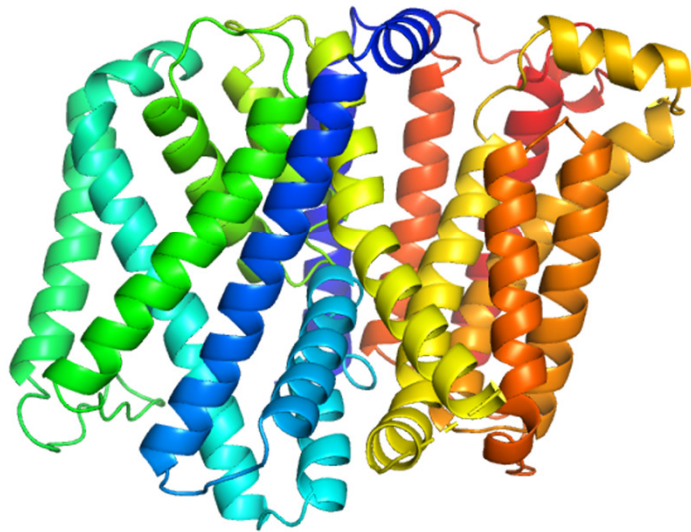
MlaE-like fold
PDBID: 7ch0
RMSD of 1.28 Å



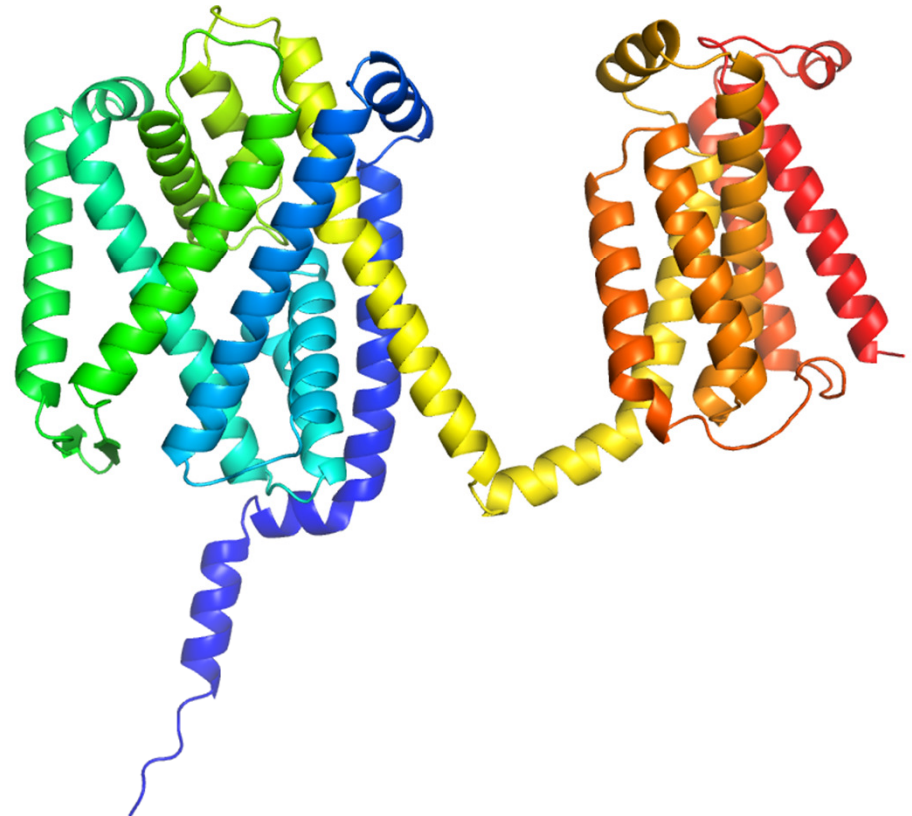
ER membrane protein complex subunit 6
PDBID: 6ww7
RMSD of 0.96 Å

Prediction of new TM folds

MprF (PDBID: 7DUW)



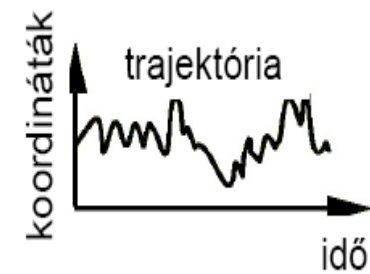
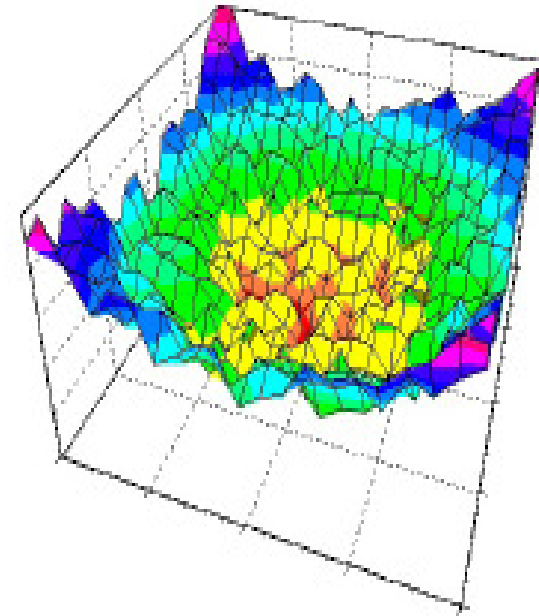
AF2



Van szerkezet, mozgassuk meg!

Molekula dinamika/Molecular dynamics

- valós potenciálfelület
- mozgásegyenletek
idő-lépésenkénti numerikus megoldása
- trajektória



A „force field”

Baker (2007)

$$E_{\text{prot}} = W_{\text{rot}} E_{\text{rot}} + W_{\text{atr}} E_{\text{atr}} + W_{\text{rep}} E_{\text{rep}} + W_{\text{solv}} E_{\text{solv}} + W_{\text{pair}} E_{\text{pair}} \\ + (W_{\text{mbenv}} E_{\text{mbenv}}) + W_{\text{hbond}} E_{\text{hbond}} - E_{\text{ref}}$$

$$E_{\text{mbenv}} = \sum_i^{\text{natom}} \Delta G_i^{\text{ref}}(z') \quad \Delta G_i^{\text{ref}}(z') = (1 - f(z')) * (\Delta G_i^{\text{ref, chex}} - \Delta G_i^{\text{ref, water}})$$

TABLE II. Atomic Contribution to the Solvation Free Energy in Water and Cyclohexane

	Water	Chex		Water	Chex
CR	-0.890	-1.350	NH3	-20.000	-1.145
CH1E	-0.187	-0.645	NC2	-10.000	-0.200
CH2E	0.372	-0.720	N	-1.000	-1.145
CH3E	1.089	-0.665	OH1	-5.920	-0.960
CR1E	0.057	-0.410	O	-5.330	-1.270
NH1	-5.950	-1.145	OC	-10.000	-0.900
NR	-3.820	-1.630	S	-3.240	-1.780
NH2	-5.450	-1.145	SH1E	-2.050	-1.855

Egyszerűsített (Coarse Grain) modellek

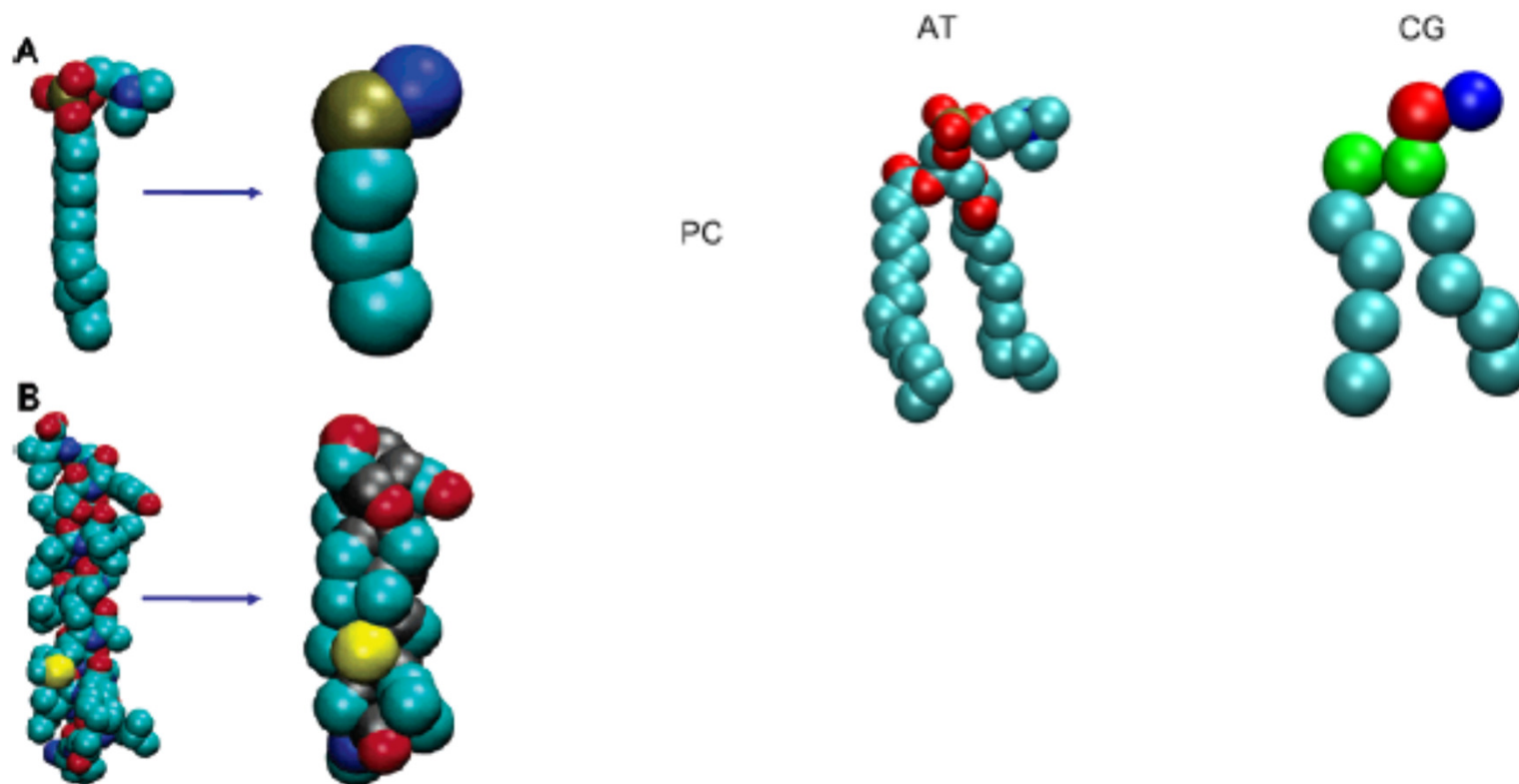
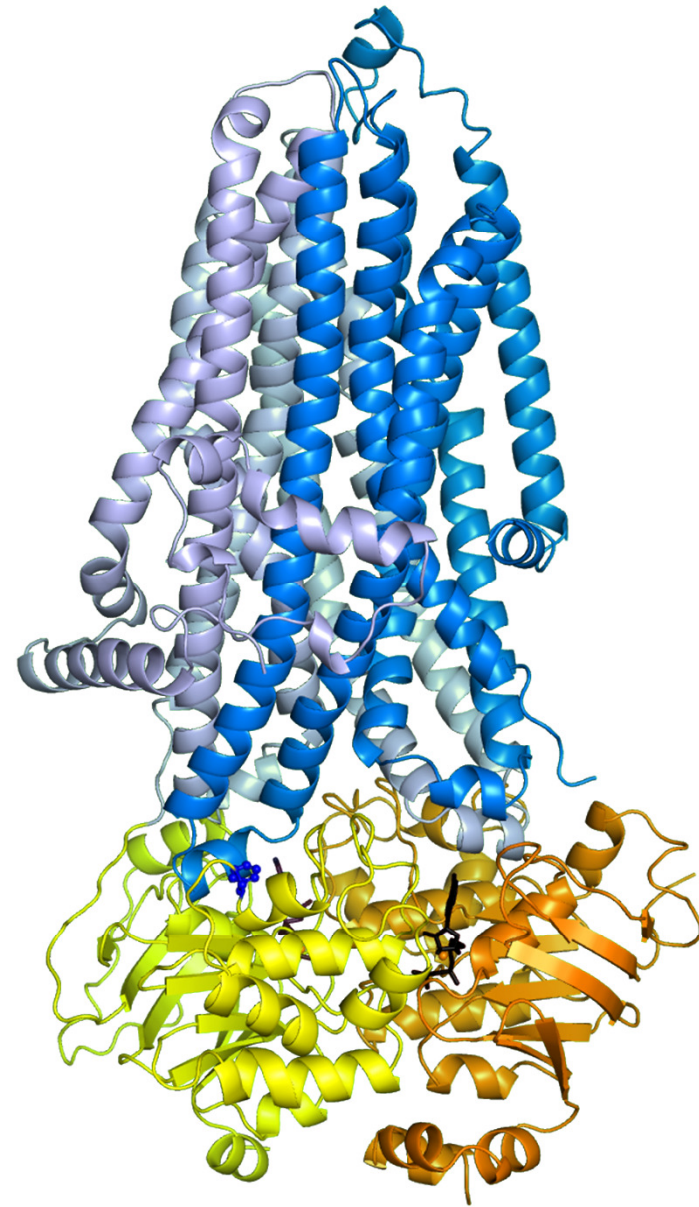
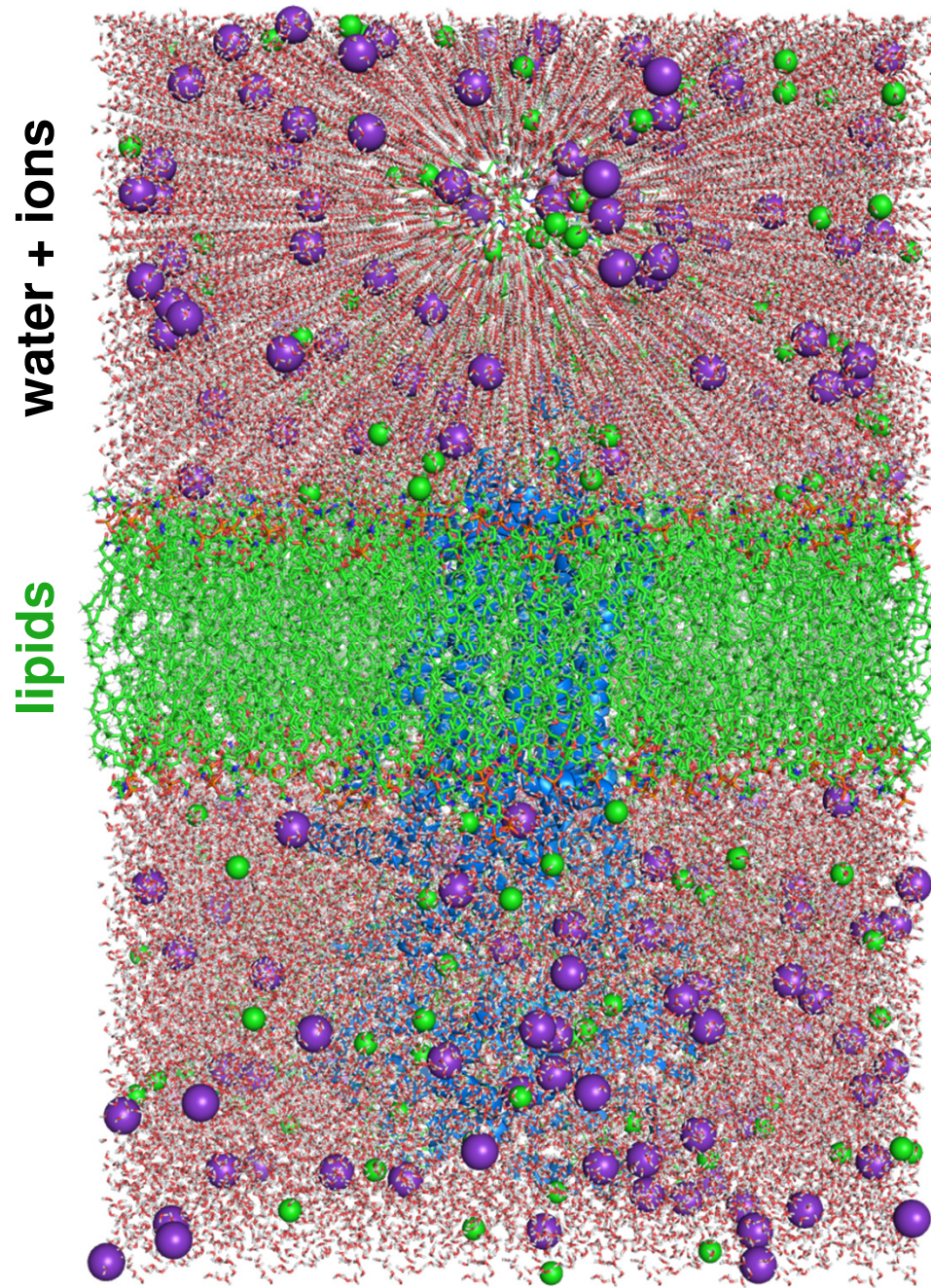


Figure 1. Atomistic (left-hand) and coarse-grained (right-hand) models compared for (A) a DPC molecule and (B) a GpA helix. Colors for atoms:

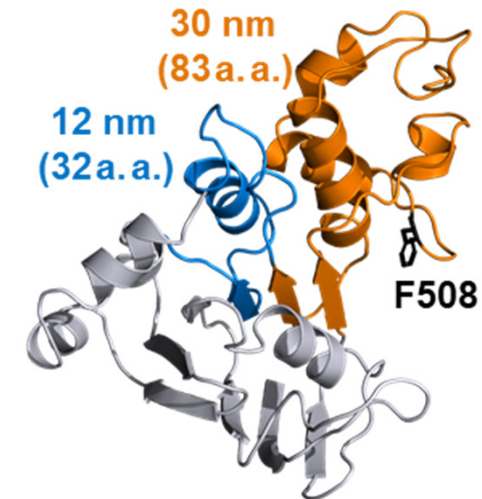
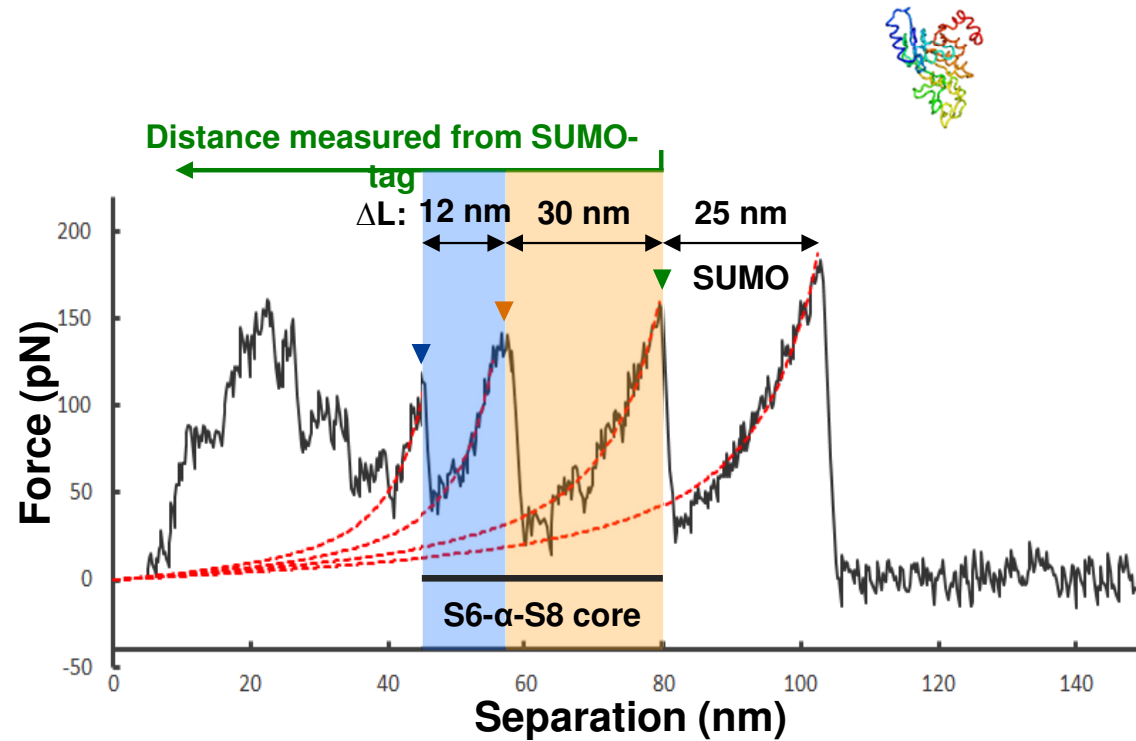
Atomistic structures



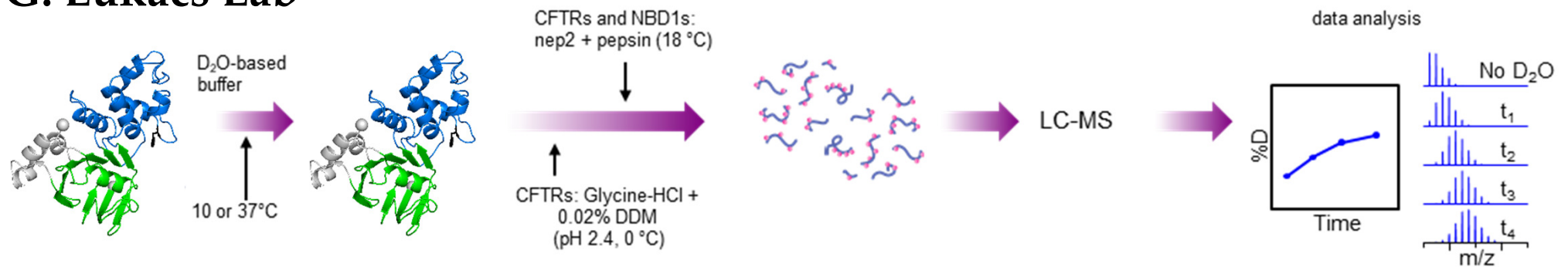
Experiments for BIA site identification

pulling with AFM, HDX

Padanyi *et al.* Comp. Struct. Biol. 2022



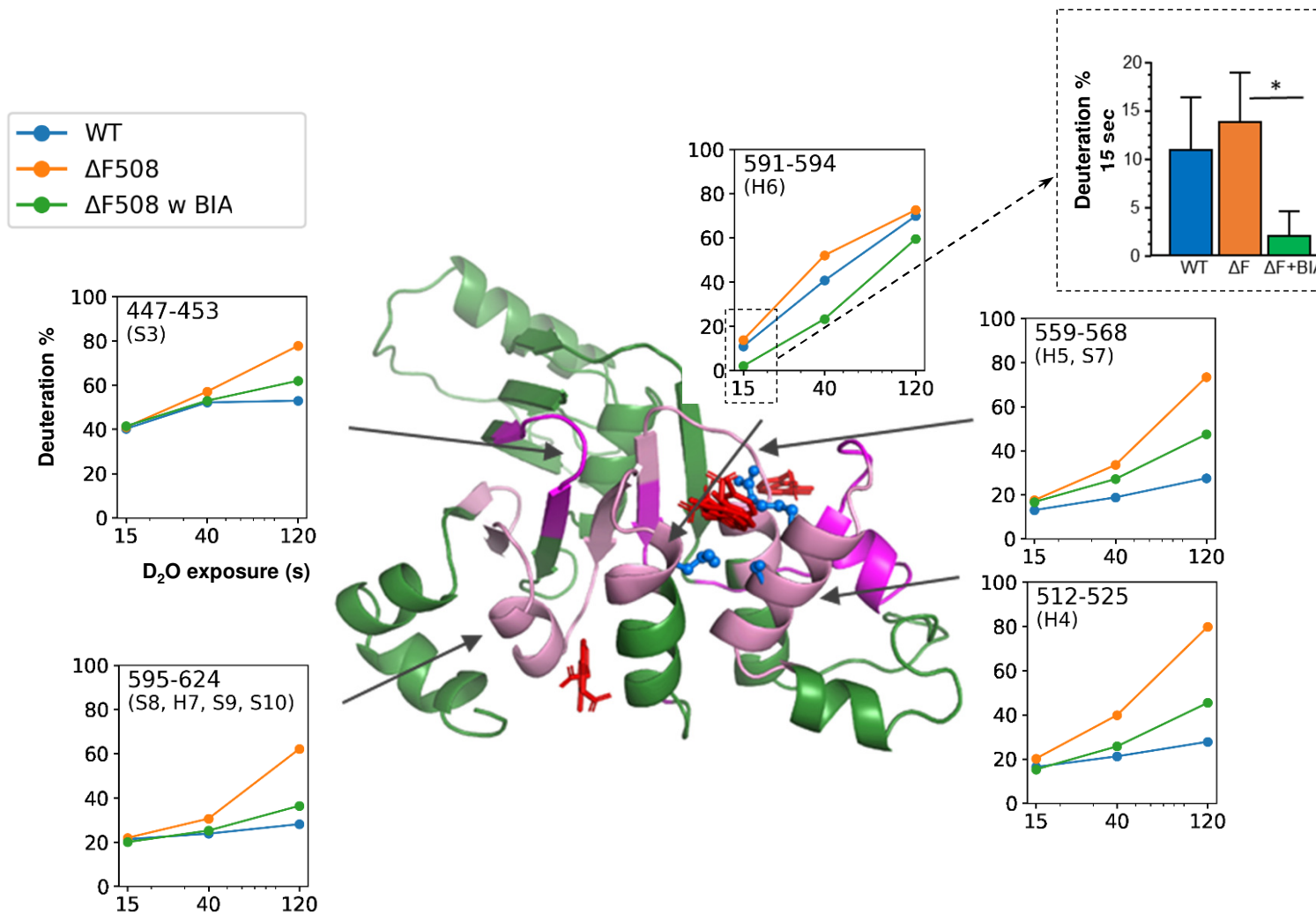
G. Lukács Lab



BIA binding site – HDX with NBD1

G. Lukács Lab

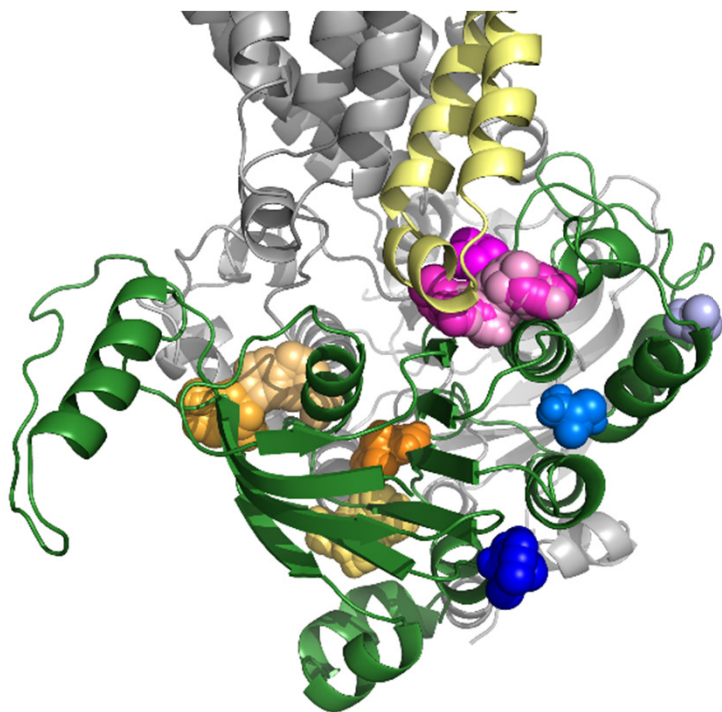
Padanyi *et al.* Comp. Struct. Biol. 2022



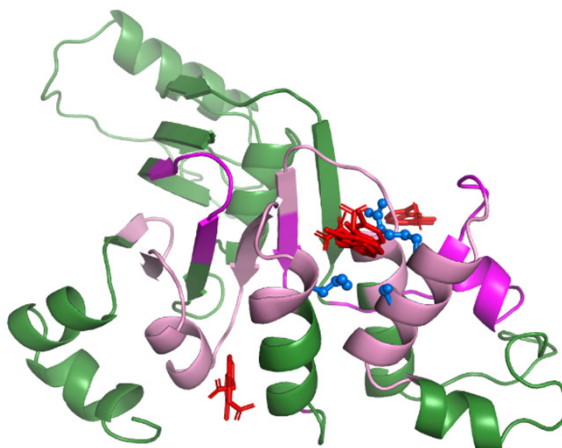
BIA binding site – *in silico*

Padanyi *et al.* Comp. Struct. Biol. 2022

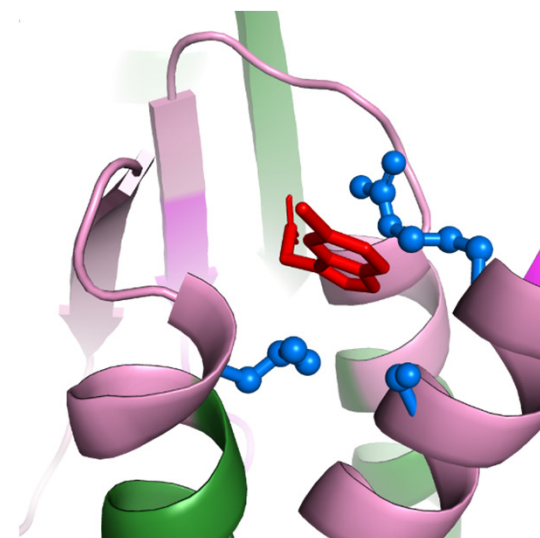
pocket detection
fpocket



docking
AutoDock Vina

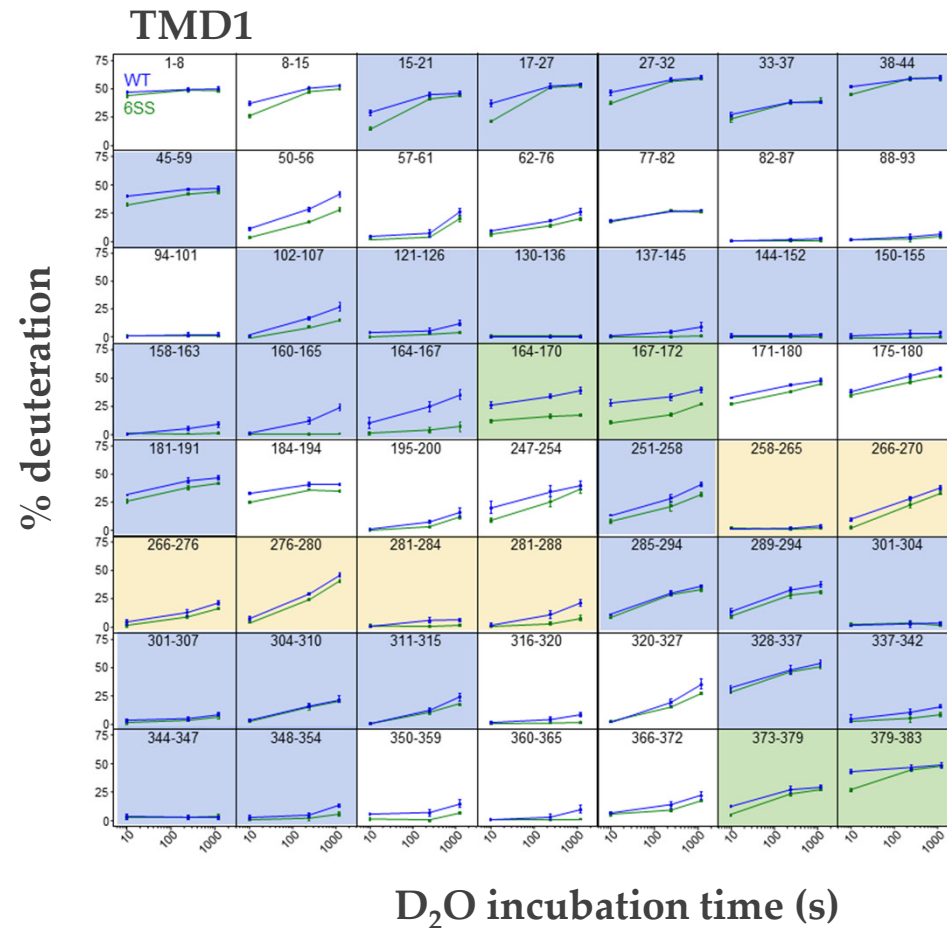
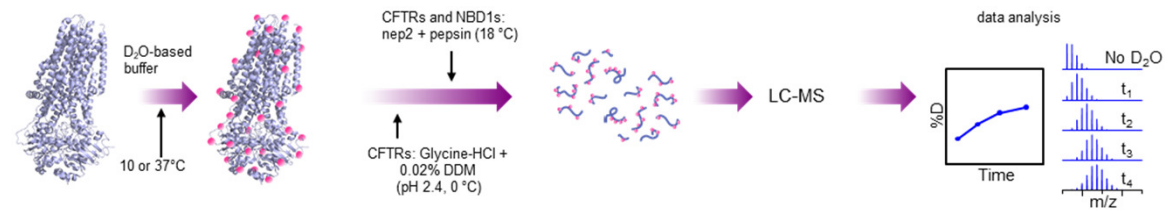


concluded site



Folding/maturation/stability of the full length CFTR

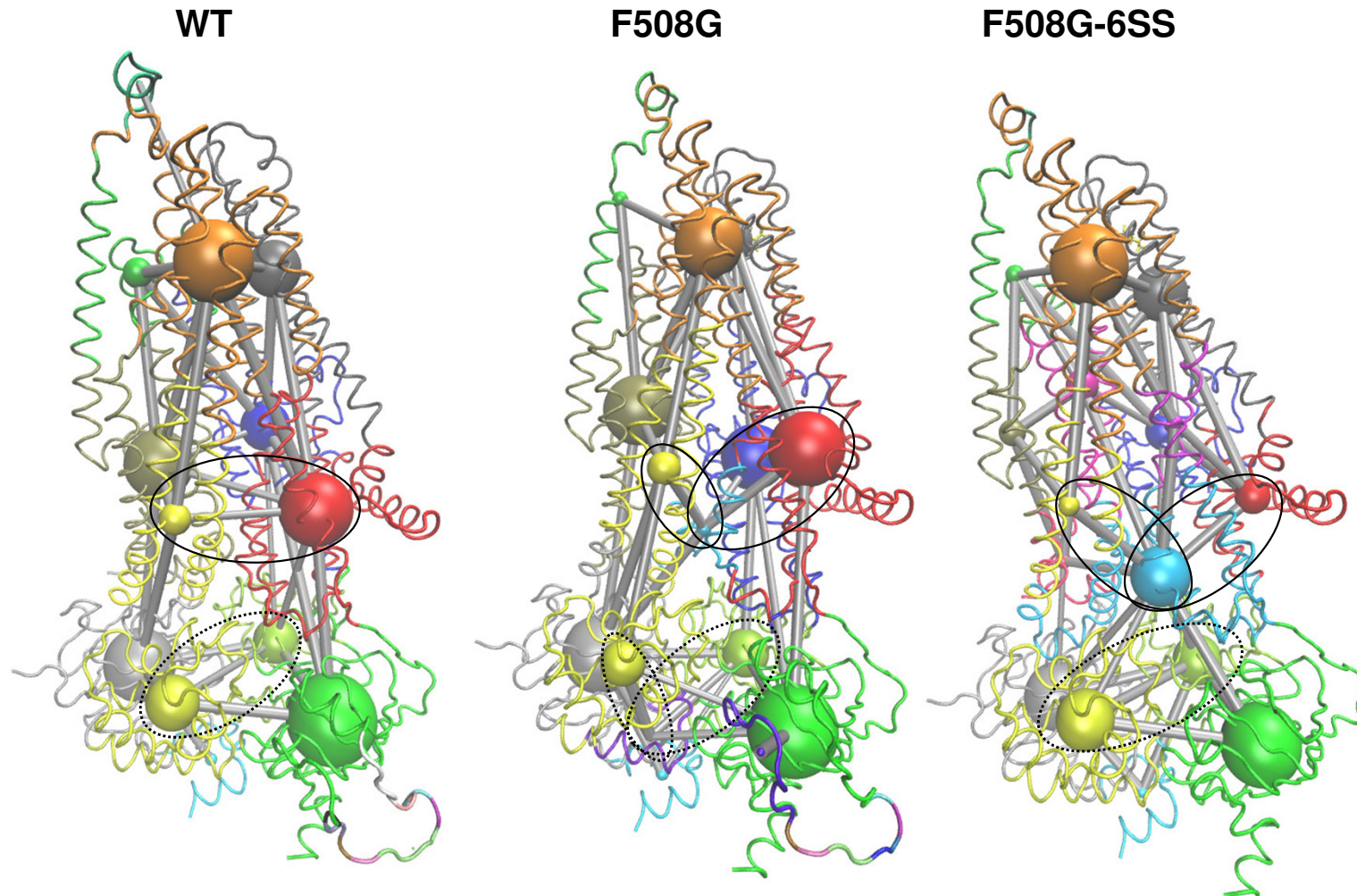
Soya *et al.* Nat. Comm. 2023



Allosteric stabilization of TM_{IC}

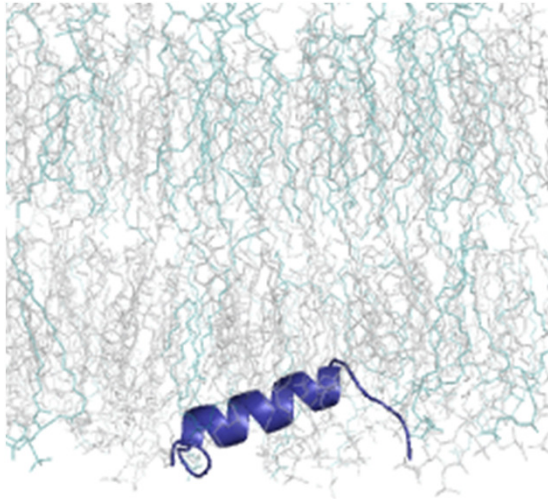
Soya *et al.* Nat. Comm. 2023

- Folded structure
- Correlated motions
- Network
- Community analysis



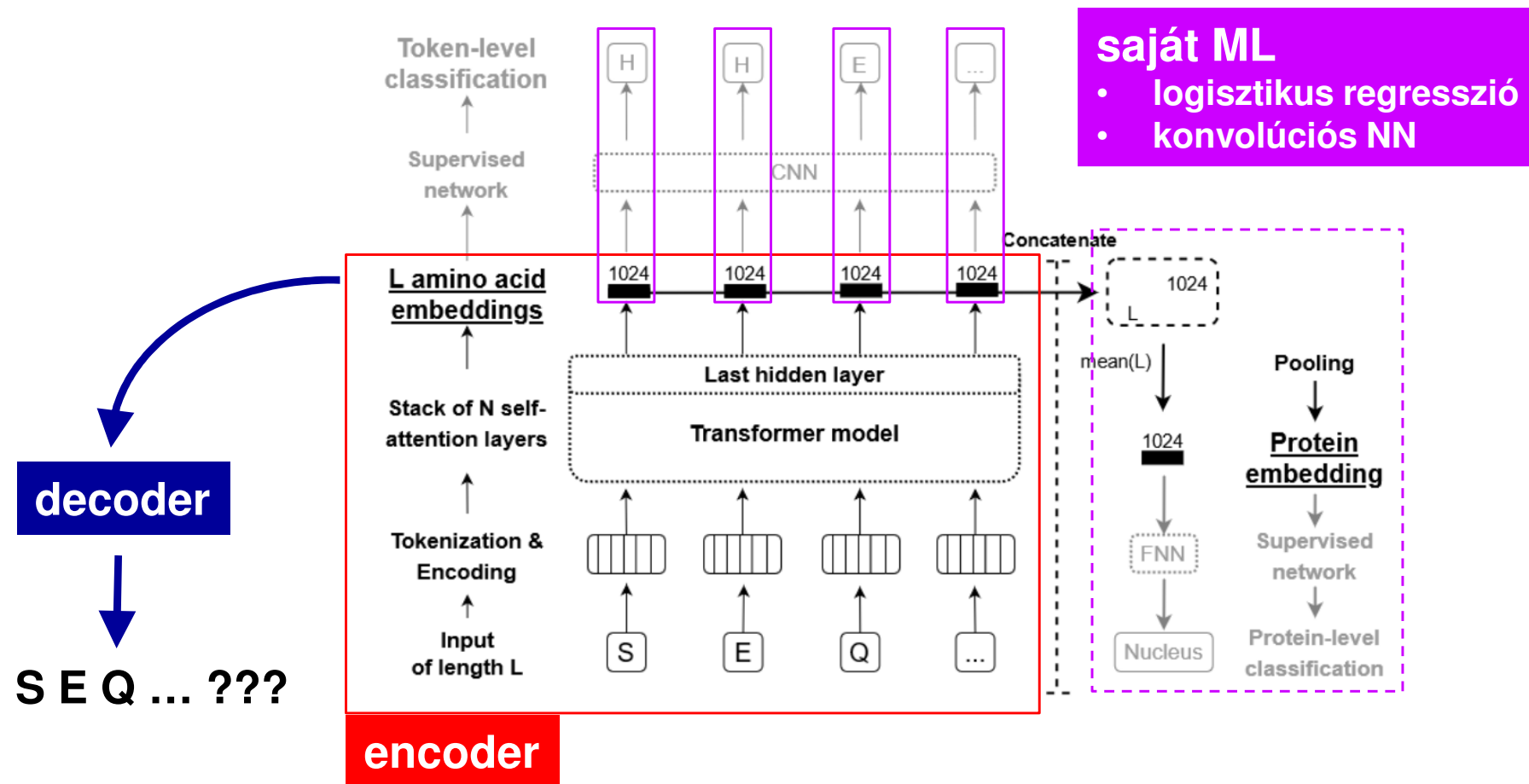
MemMoRFs

Membrane Molecular Recognition Features, <https://memmorf.hegelab.org>

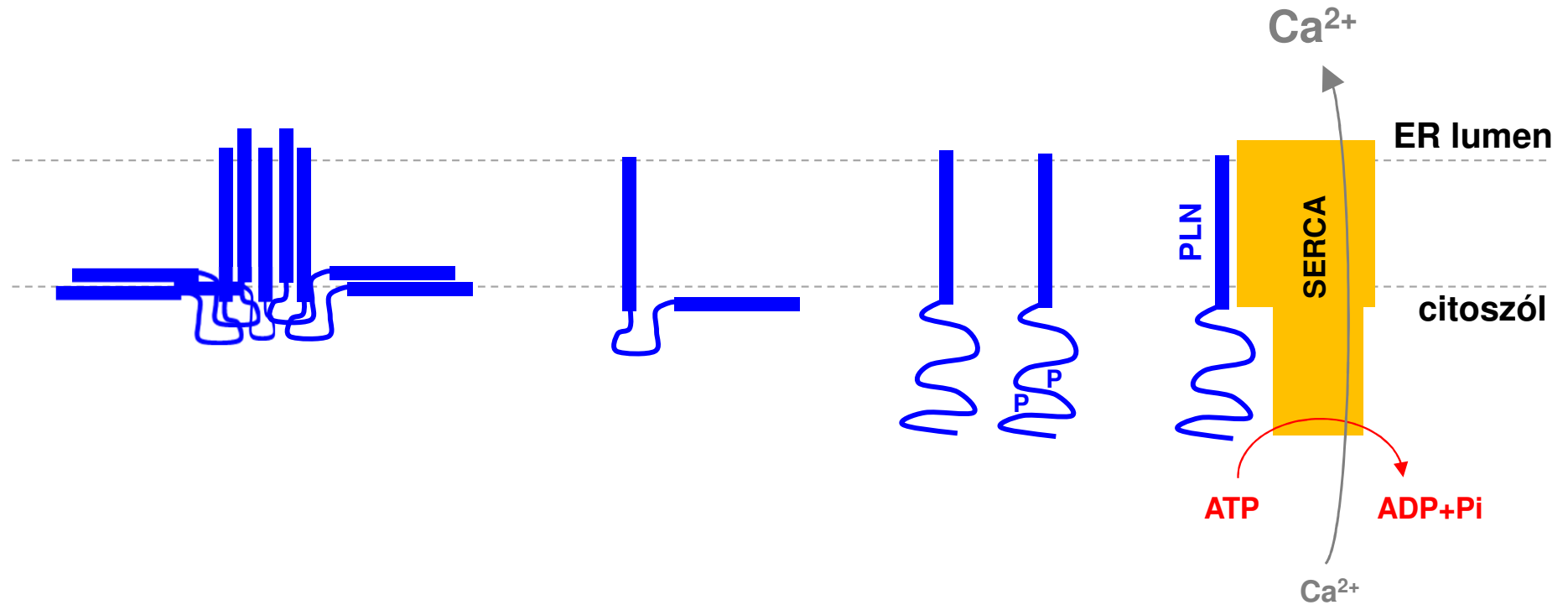


MemMoRF prediktor

protein Language Models



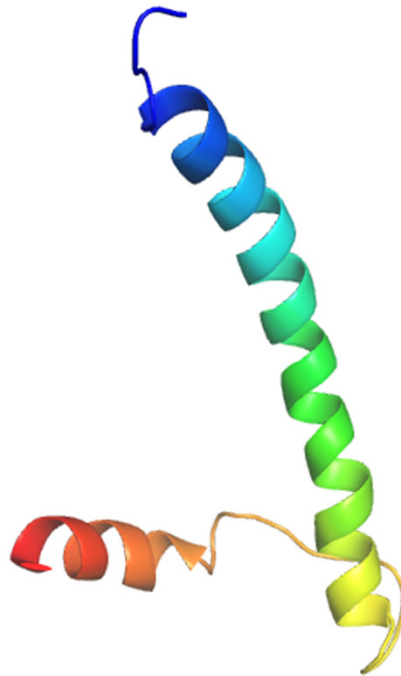
Kalcium ATP-ázok (SERCA, PMCA)



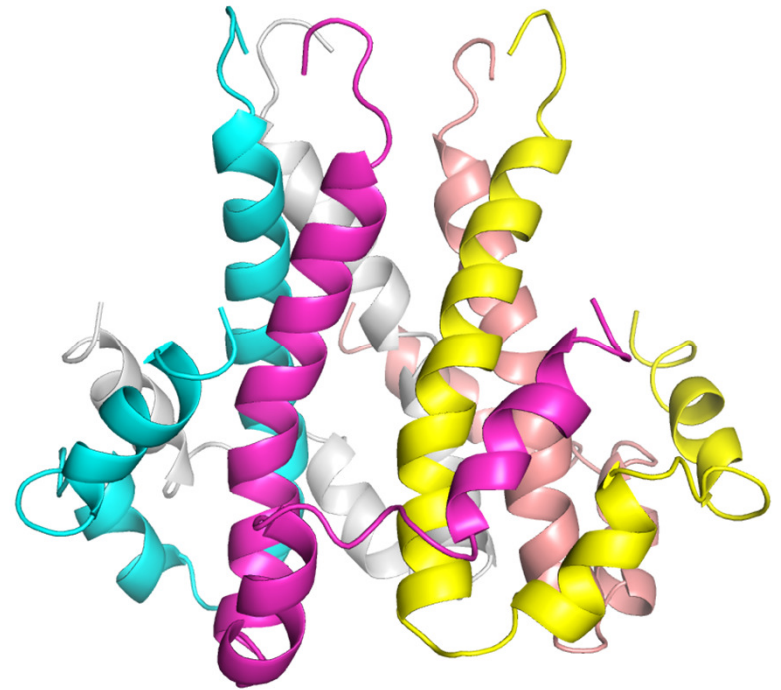
Regulinok és exoregulinok

SLN
ELN
ALN
Dworf
???

SARS Cov-2
Envelope (E) protein

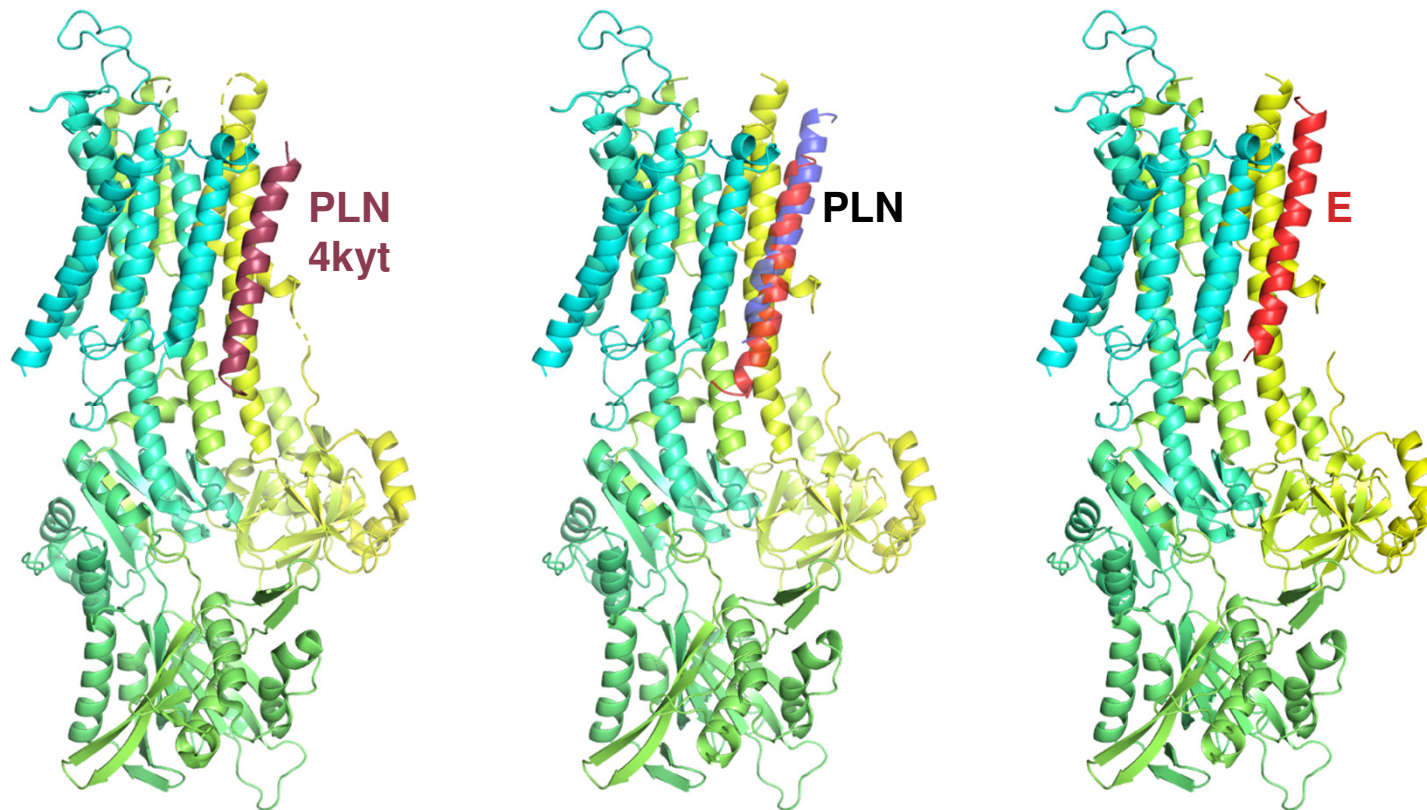


- ERGIC
- viroporin
- PDZ-kötés
- egyéb PPI?



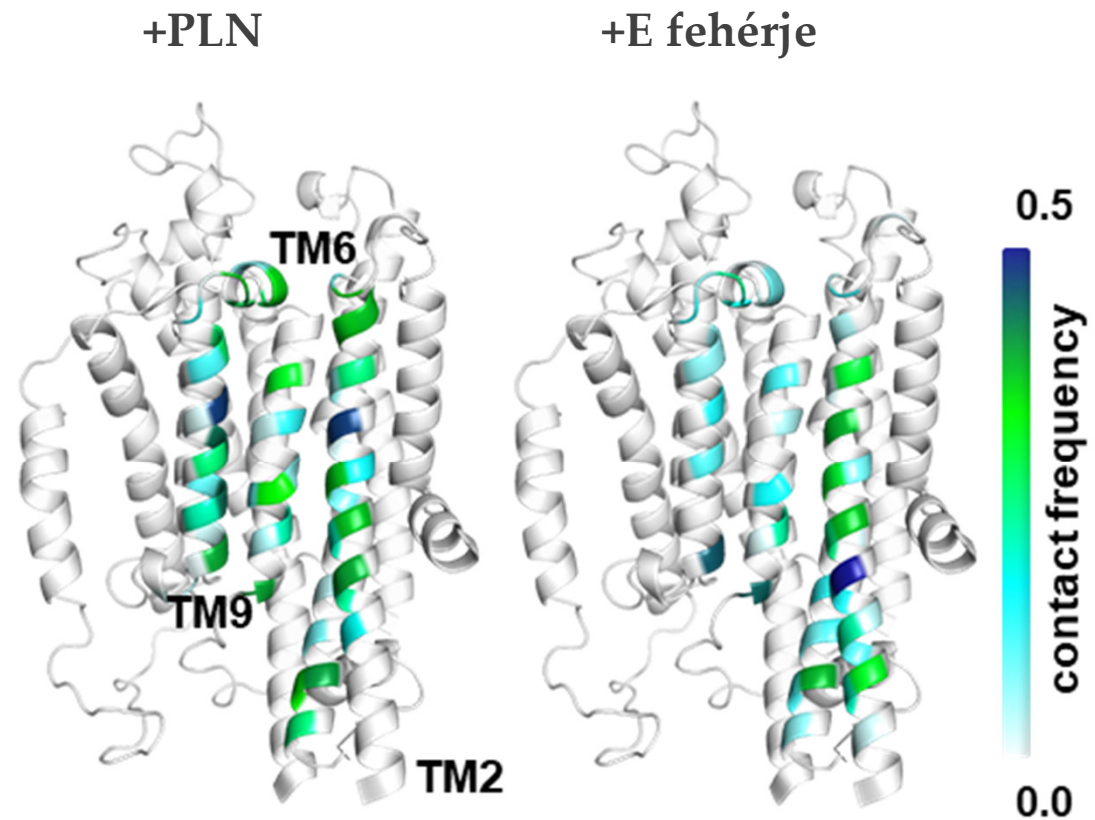
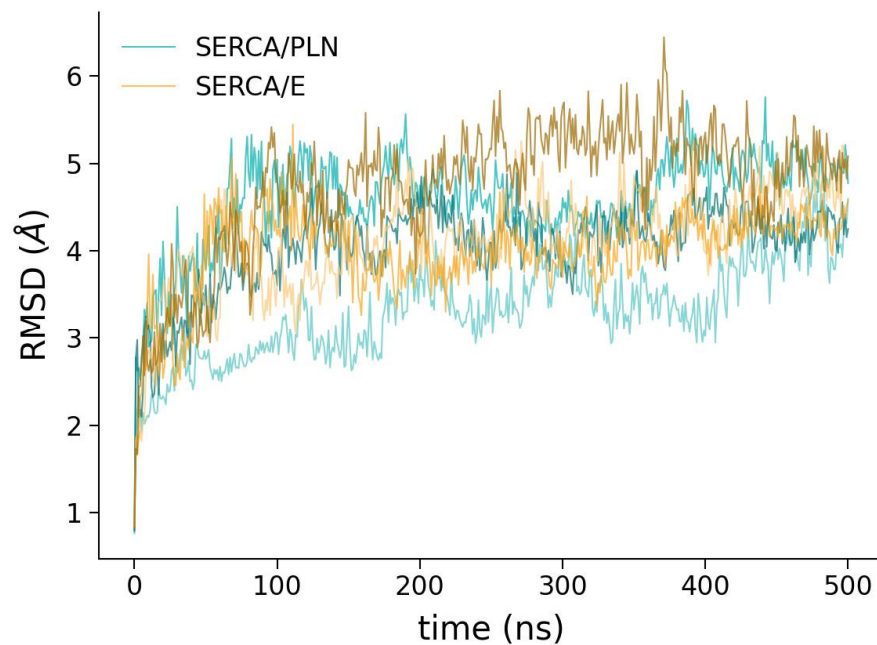
A SERCA / E protein komplex szerkezete

AlphaFold-Multimer
PIPER/ClusPro



SERCA/regulin komplexek stabilitása

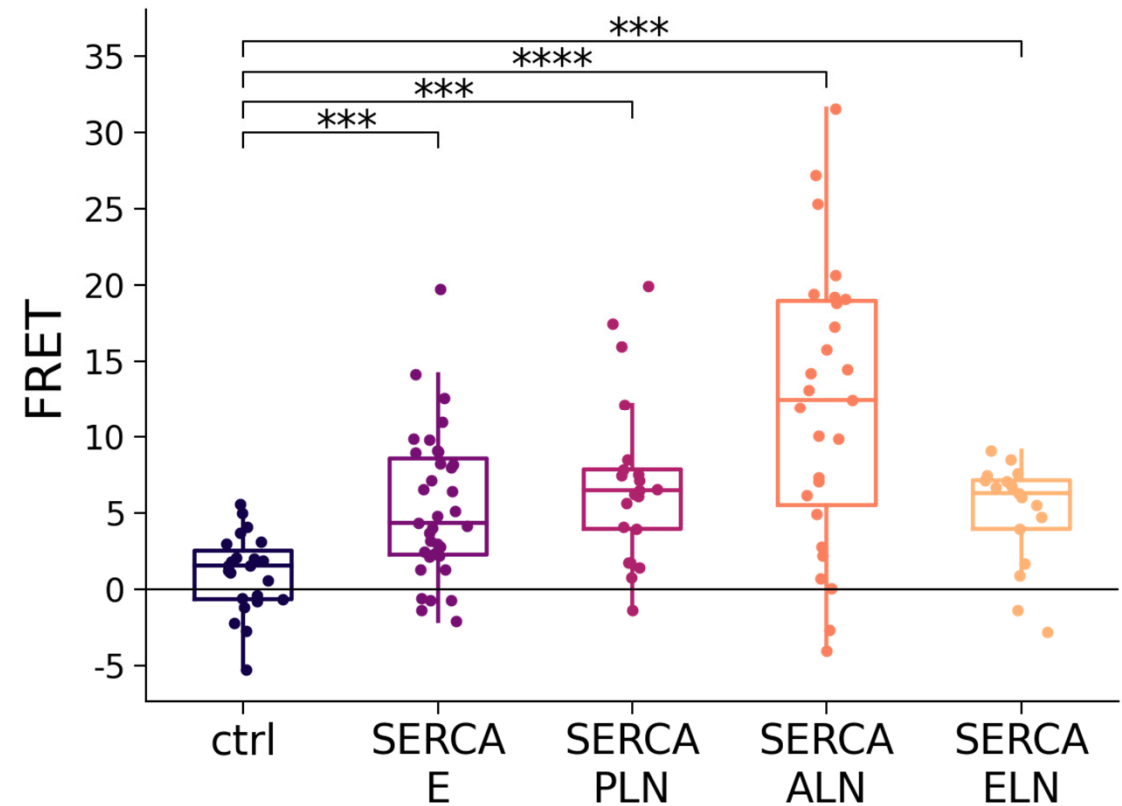
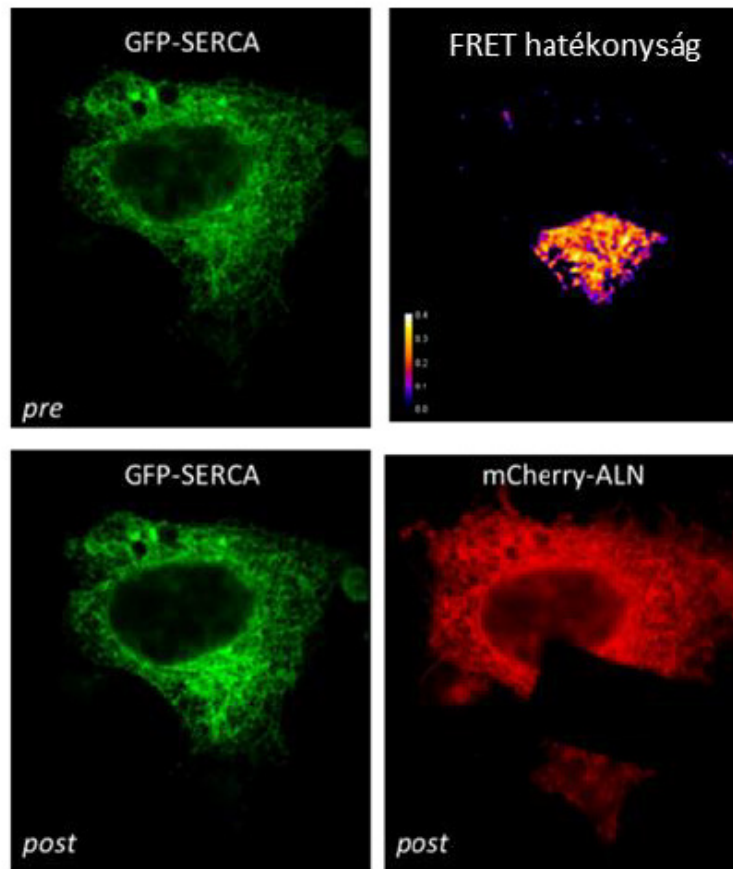
MD szimulációk



Az E fehérje kölcsönhat a SERCA pumpával

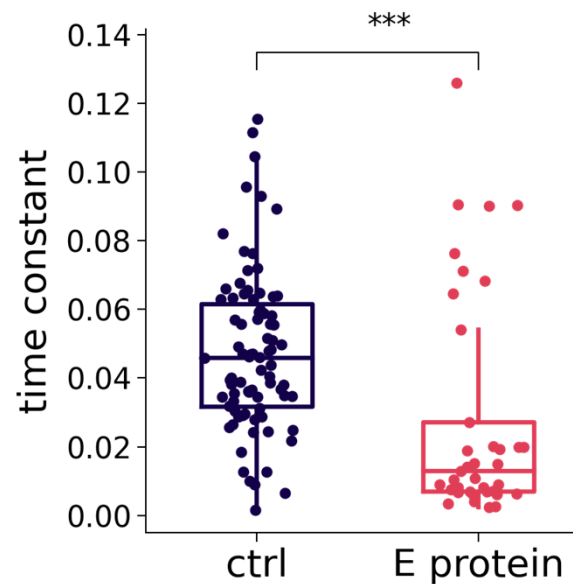
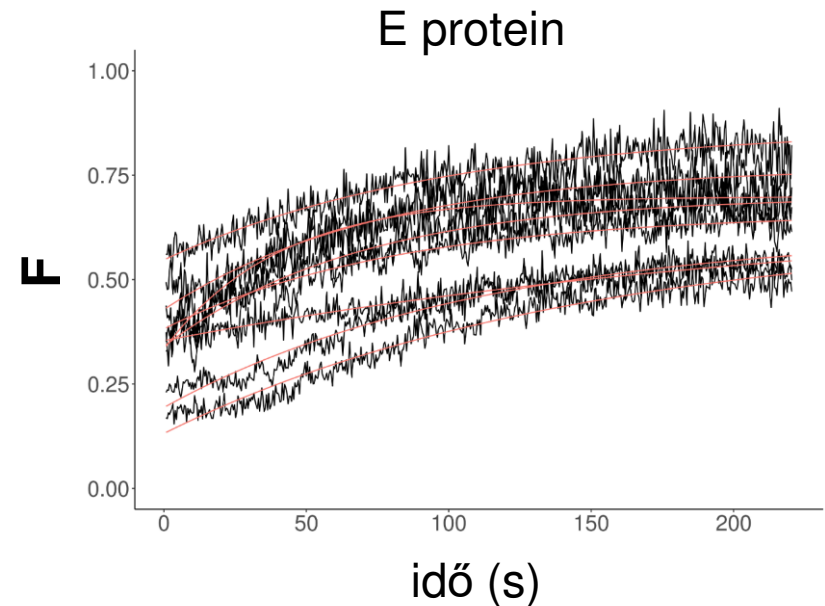
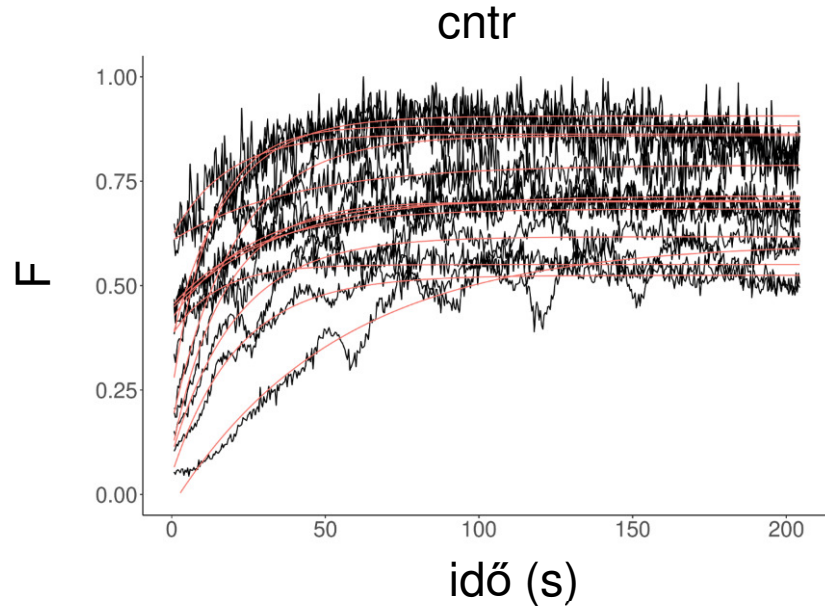
- GFP és mCherry jelölt fehérjék
- koexpresszió HeLa sejtekben
- Lokalizáció átfedés
- AP-FRET

PR: P45



Az E fehérje gátolja a SERCA működését

nincs EC Ca^{2+}
ER ürítése
+ Ca^{2+}



AlphaMissense.hegelab.org

1,600 users in March

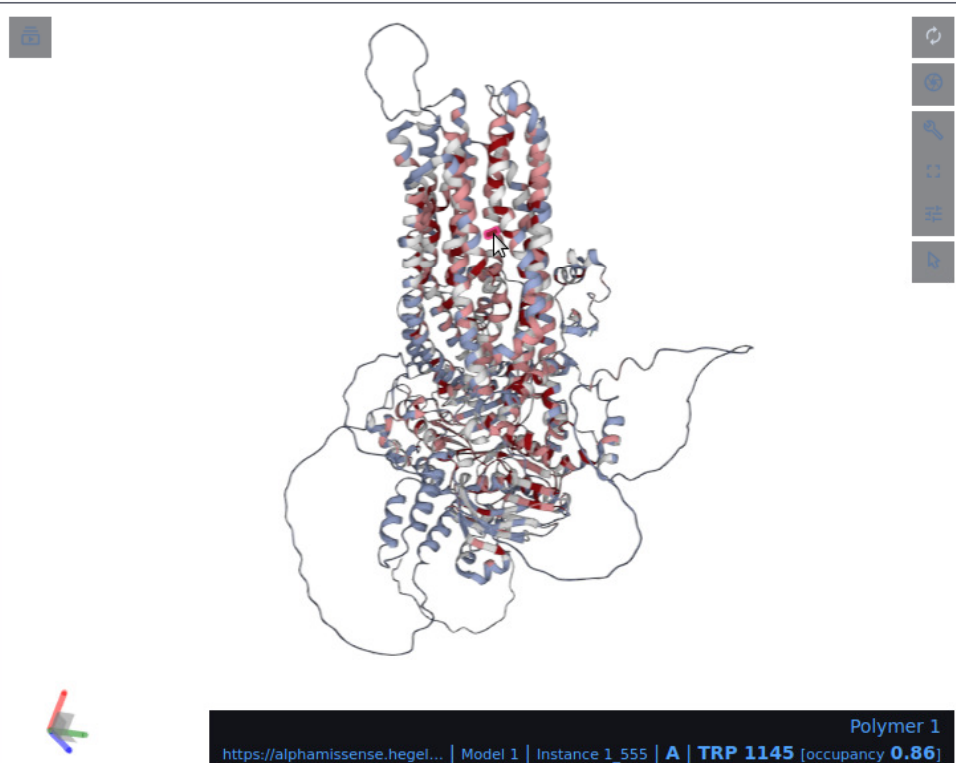
>>> Google DeepMind contacted us

[Home](#) [Search](#) [Hotspot](#) [Download](#) [Help](#)

Spot mutations

Identifier: [Get](#)

CFTR (CFTR_HUMAN, P13569, ENST00000003084.10) [Download](#)



Polymer 1
<https://alphamissense.hegelab.org> | Model 1 | Instance 1_555 | **A** | **TRP 1145** [occupancy **0.86**]

a.a.	benign	ambiguous	pathogenic	mean
Q2	5:K,R,L,P,E	1:H		0.268
R3	3:T,K,G	3:S,W,M		0.338
S4	2:T,A		3:L,W,P	0.492
P5		1:A	5:T,R,L,S,H	0.603
L6	4:V,M,Q,R	1:P		0.237
E7	3:D,Q,G	2:V,A	1:K	0.403
K8	6:T,R,N,Q,E,M			0.156
A9	1:S	1:G	4:D,V,P,T	0.622
S10	4:N,T,C,G	2:I,R		0.217
V11	6:I,F,L,D,A,G			0.139
V12	5:I,F,L,A,G		1:D	0.189
S13	1:A	2:T,C	3:F,P,Y	0.525
K14	5:I,T,R,Q,E	1:N		0.265
L15	2:E,I,V		3:D,H,P	0.286

Összefoglalás

- Fehérje szerkezet
- Fehérje dinamika
- Számításos módszerek igen fontosak
- Fehérjék működésének és mutációk hatásának megismerése
- Gyógyszertervezés

www.hegelab.org
hegedus.tamas@hegelab.org