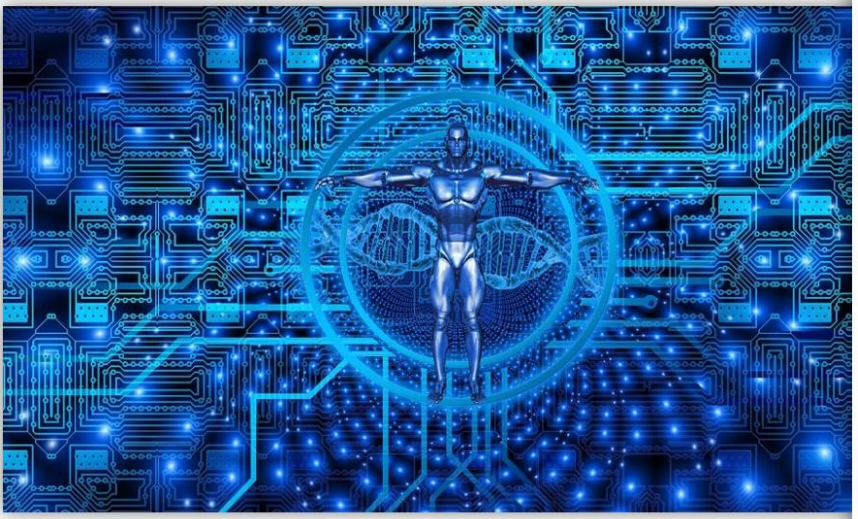# Structure-based drug discovery case study: CFTR modulators

Pasqualina Dursi
Institute of Advanced Biomedical Technologies
National Research Council

The Bioinformatics Lab consists of approximately 15 researchers and technologists having a diverse background encompassing bioinformatics, biotechnology, engineering, physics and computer science.

The Lab develops, implements and applies advanced analysis techniques in order to tackle open problems from the Life Sciences realm, with particular attention to key sectors such as genomics, epigenomics, bulk and single-cell transcriptomics, multi-omics data integration, molecular biodiversity, systems biology, molecular modelling and artificial intelligence.

Additionally, biological data banks and bioinformatics tools are designed and developed within the Lab for Big Data analyses in the sectors of biomedicine, nutrition and environmental problems. A Translational Bioinformatics unit also part of this Lab provides preparation of biological samples and performs related *omics analyses so to further support and validate the bioinformatics predictions.

The Lab has built and manages an advanced HPC (High Performance Computing) infrastructure so to support the Lab projects' computational needs from the inside. The infrastructure provides predictable execution times for the computations, without

STAFF

Federica Chiappori

Researcher

Francesca Cupaioli

Researcher

Pasqualina D'Ursi

Researcher

Matteo Gnocchi

Technologist

Andrea Manconi

Researcher

Ivan Merelli

Researcher

Alessandra Mezzelani

Researcher

Ettore Mosca

Researcher

Marco Moscatelli

Technologist

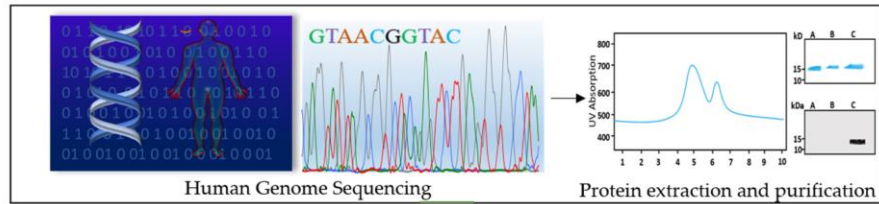**Alessandro Orro**
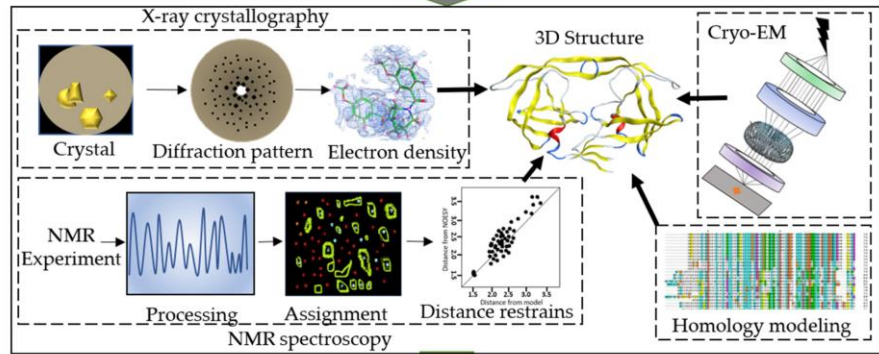
Researcher

Matteo Uggeri
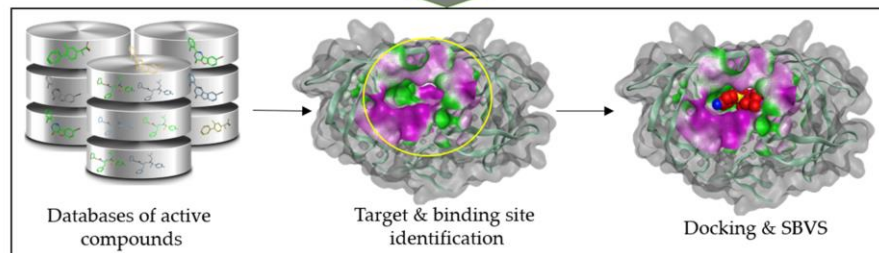
John Hatton

Luciano

# A workflow diagram of structure-based drug discovery process

Human genome sequencing followed by extraction and purification of the target proteins

Structure determination of the therapeutically important proteins using integrative structural biology approaches.
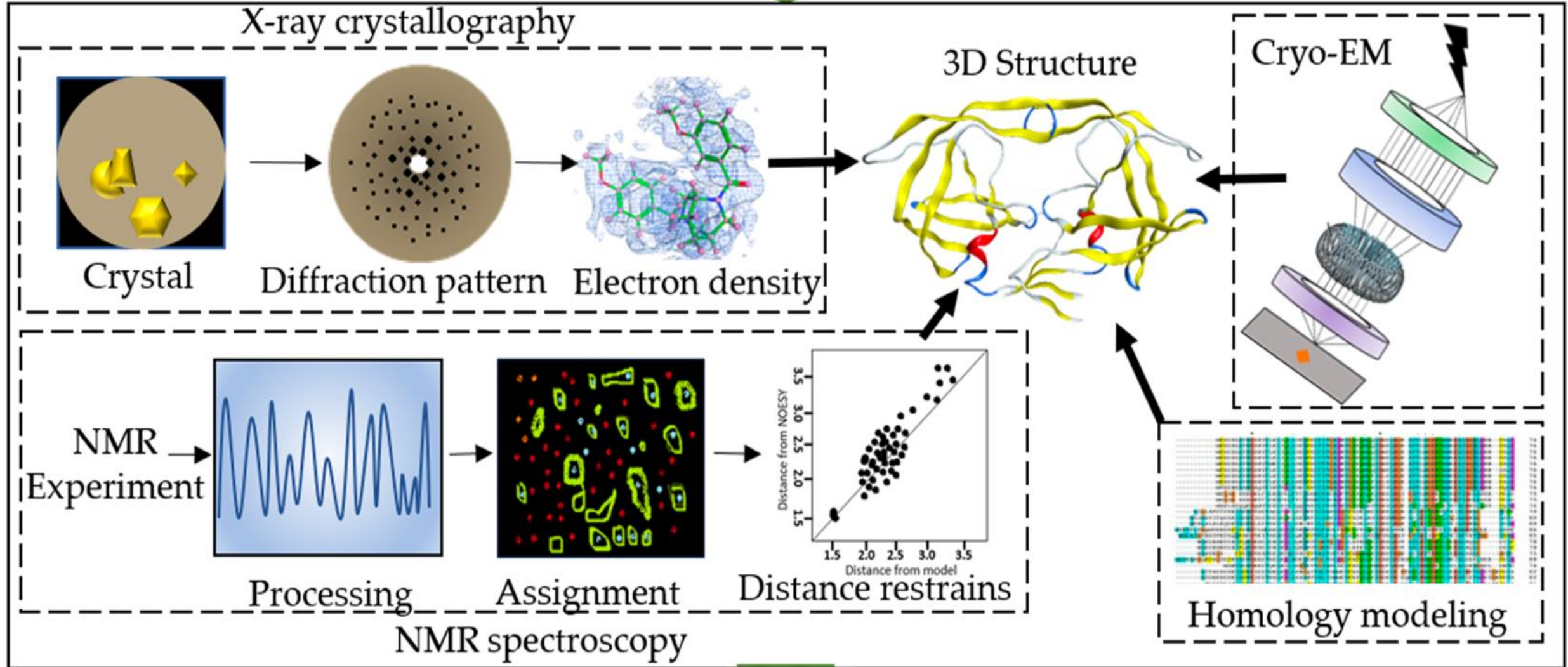
Database preparation of the active compounds.
Identification of the druggable target protein and its binding site.
The databases of active compounds are screened and docked into the binding cavity of the target protein.
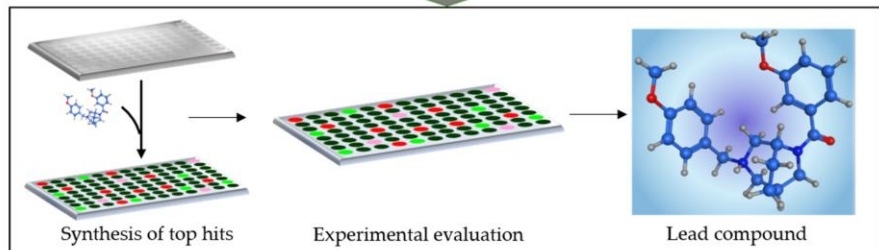
The top hit compounds obtained as a result of virtual screening and docking are synthesized and tested in vitro. Further modifications can be done for the optimization of the lead compound.

Structure-based drug discovery

3D Structure of protein target

YouTube   Maps

RCSB PDB   Deposit ▾   Search ▾   Visualize ▾   Analyze ▾   Download ▾   Learn ▾   About ▾   Documentation ▾   Careers   COVID-19   MyPDB ▾   Contact us

RCSB PDB   PROTEIN DATA BANK

209,389 Structures from the PDB
1,068,577 Computed Structure Models (CSM)

▾ 3D Structures ⓘ   ]]   Include CSM ⓘ ◯   🔍

Advanced Search | Browse Annotations   Help

PDB-101   wwPDB   EMDataResource   NAKB   wwPDB Foundation   PDB-Dev

💻 New: More Computed Structure Models (CSM) available   Learn more

- 🔖 Welcome
- ☁ Deposit
- 🔍 Search
- 🖼 Visualize
- ▦ Analyze
- ☁ Download
- 📖 Learn

RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

Experimentally-determined 3D structures from the **Protein Data Bank (PDB)** archive

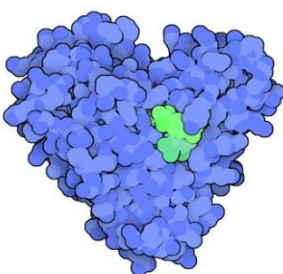**Computed Structure Models (CSM)** from AlphaFold DB and ModelArchive

These data can be explored in context of external annotations providing a structural view of biology.
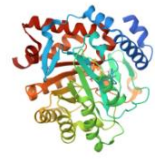
Explore NEW Features   PDB-101 Training Resources

September **Molecule of the Month**

Histone Deacetylases

Latest Entries
As of Tue Sep 05 2023

8OFW

Crystal structure of the full-length dihydroorotate dehydrogenase from Mycobacterium tuberculosis

Features & Highlights

Take the Tabular Reports Survey and Win
Please take this brief survey about RCSB.org tabular reports to be entered into a drawing for a set of Bound Playing Cards.

Turning Global Data into Global Knowledge
RCSB PDB is highlighted in *EdgeDiscovery*, a publication that highlights research and research computing initiatives in New Jersey and beyond

Updated Annotation and Standardization of Peptide Residues
In October 2023, wwPDB will roll out updated CCD data files with

News   Publications ▾

Head Back to School with PDB-101
Use PDB-101 resources, training materials, and curricula to explore biology from AAA+ Proteases to Zika Virus
» 09/05/2023

Education Corner: Empowering Educators
Learn how the Schrödinger Education team describes *Empowering Educators with Research-Grade Computational Tools*.
» 08/29/2023

Create Videos About Structural Science
All members of the American

📊 **PDB at a Glance**   **64,309** Structures of Human Sequences   **16,612** Nucleic Acid Containing Structures

📊 **CSM at a Glance**   **999,251** AlphaFoldDB   **69,326** ModelArchive

More Statistics

About   About Us

Help   Contact Us

RCSB PDB (citation) is hosted by

RCSB PDB is a member of the

RCSB Partners
Nucleic Acid Knowledgebase
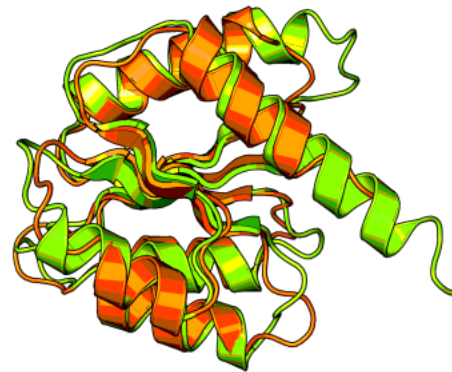
Structure prediction methods

Homology modelling

The principle behind it is if two proteins share high sequence similarity.

LUC-A
LUCIA

sequence protein A `MNSFSTSAFGPVAFSLGLLLVLPAAFP-APVPPGEDSKDVAAPHRQPLTS`
`|...|...|.||||  |||:||...|||  :.|..|:.::|  ..|:|  |:.:`
sequence protein B `MKFLSARDFHPVAF-LGLMLVTTTAFPTSQVRRGDFTED-TTPNR-PVYT`

sequence protein A `SERIDKQIRYILDGISALRKETCNKSNMCESSKEALAENNLNLPKMAEKD`
`:.:;...|.::|..|..:|||.||.::.|.::.:|||||||.||::..|`
sequence protein B `TSQVGGLITHVLWEIVEMRKELCNGNSDCMNNDDALAENNLKLPEIQRND`

They have similar 3D structure

Protein A and protein B are coloured in green and orange respectively and their 3D structures are superimposed

Structure prediction methods

Homology modelling

How is a 3D model created?

Template   **Unknown structure**

ELAGIILTVSYIPSAEKIA   ELAIGILTVSYIPSAEKIR

Sequence alignment

ELAGI-ILGVSYIPSAEKI-ARACELTI
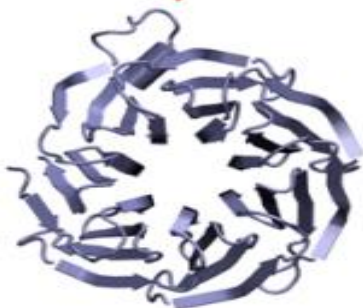ELA-IGILTVSYIPSAEKIRAP--ELTI

Structural model

If one of the protein sequences has a known structure, then the structure can be copied to the unknown protein and the structural model is obtained

Structure prediction methods

Threading
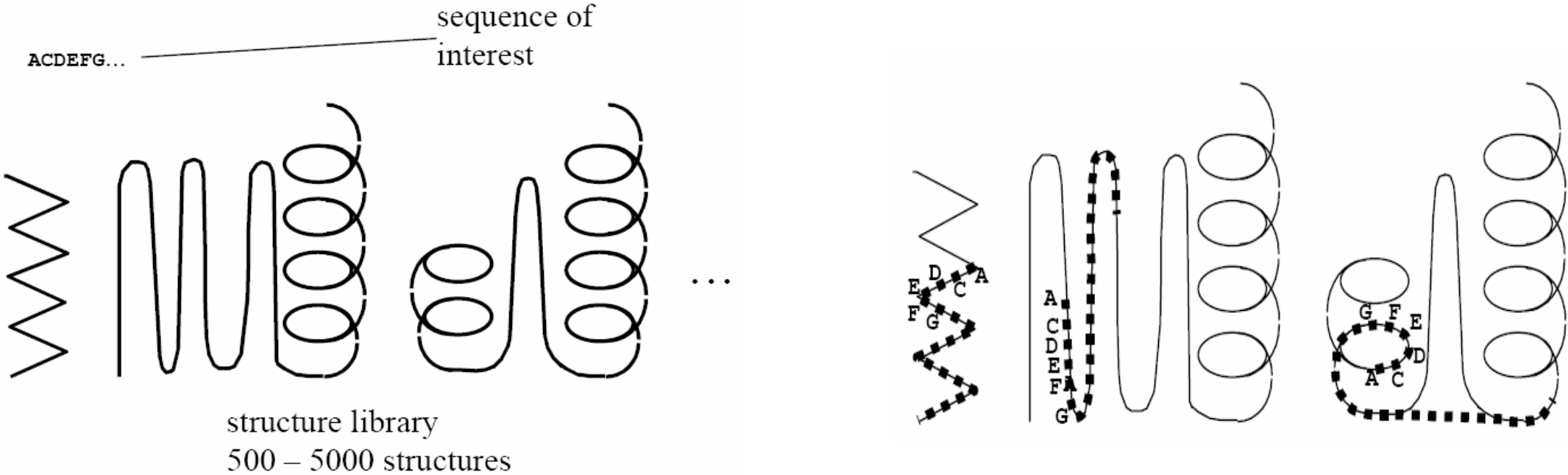
Threading takes a protein sequence and a library of templates or known structures

The sequence ACDEFG is dragged through each location on each template, and then the best matching structural fold is found using energy-based criteria.
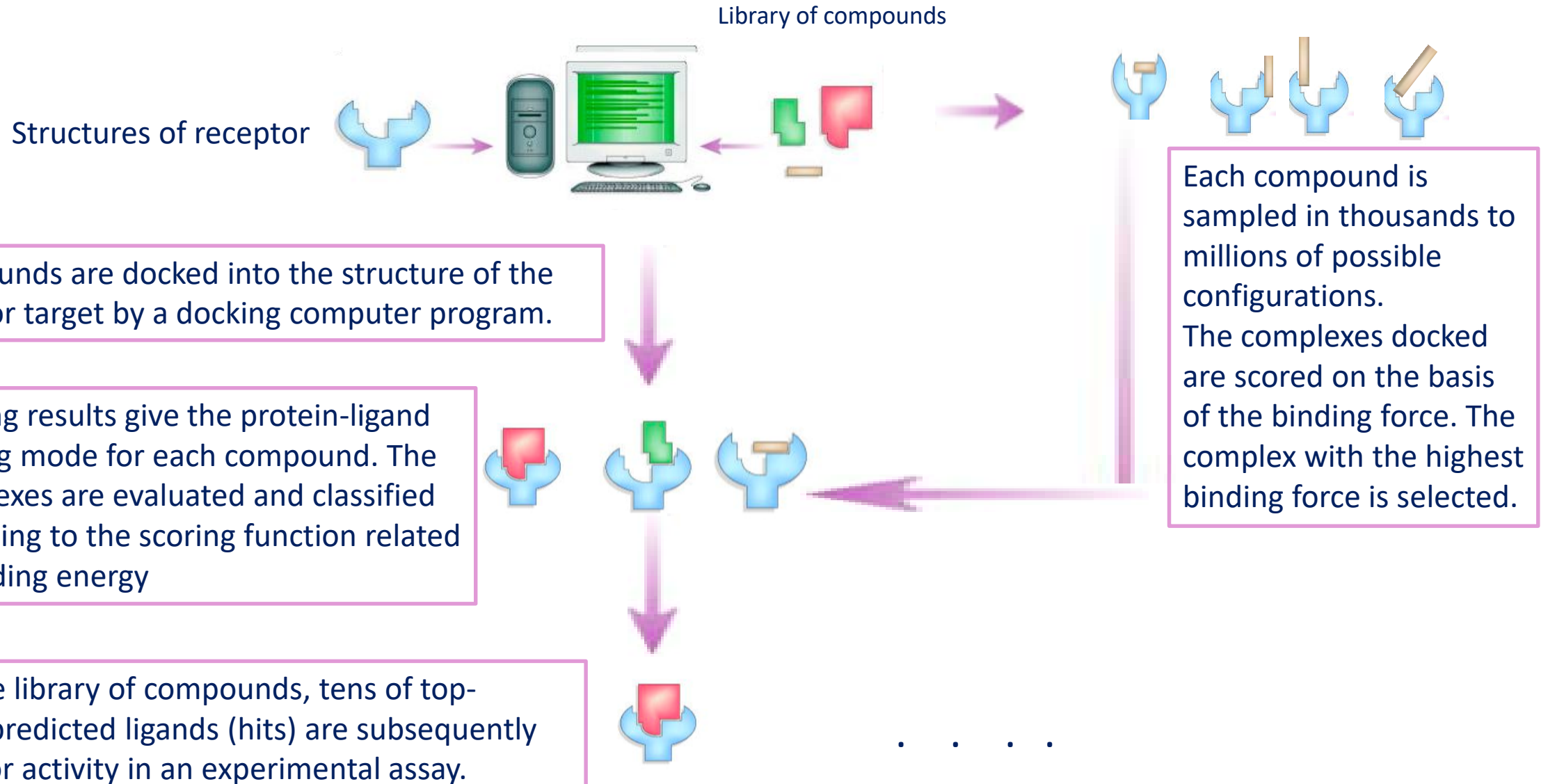
Binding pocket identification

Table 15.1. Available software for pocket identification, binding site comparison and chemical compound database search.

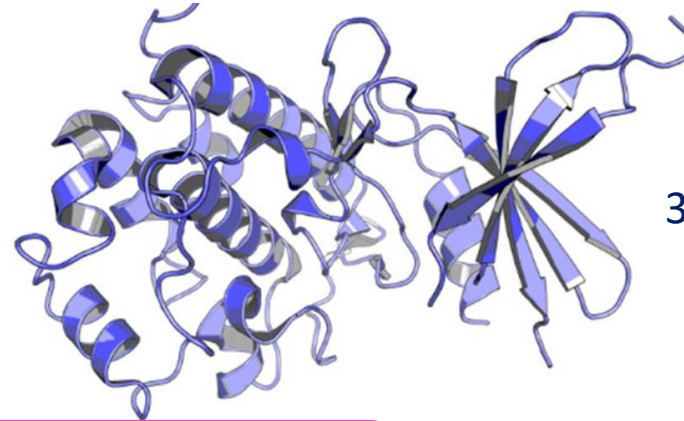| Category	Software	Feature	Web address
Pocket identification	SURFNET	Identifies a pocket with a large and a small probe sphere rolling on the protein surface	www.ebi.ac.uk/thornton-srv/software/SURFNET [104]
	LIGSITEcsc	Identifies a ligand-binding pocket as a spatially semiclosed region by protein surface and shows sequence conservation	http://projects.biotec.tu-dresden.de/pocket [105]
	CAST	Identifies a pocket using the Delaunay triangulation and the α complex	http://sts.bioengr.uic.edu/castp [106]
	PASS	Identifies a pocket as a surface region where small probe spheres fit into	www.ccl.net/cca/software/UNIX/pass/overview.html [107]
	VisGrid	Identifies a pocket as a surface position with a low visibility	http://kiharalab.org/VisGrid [101]
Binding site comparison	eF-Site	Compares shape and the surface electrostatic potential using a graph matching method	http://ef-site.hgc.jp [108]
	SitesBase	Compares atoms in pockets using geometric hashing	www.modelling.leeds.ac.uk/sb [109]
	Pocket-Surfer	3DZD-based method described in this chapter	http://kiharalab.org/pocket-surfer [103]
Chemical compound database search	SIMCOMP	Subgraph matching method	www.genome.jp/tools/simcomp [110]
	UNITY2D	Fingerprint-based method, a molecule is represented as a Boolean array of features of the molecule	A part of SYBYL package www.tripos.com [111]
	PubChem	Users can draw query molecules or specify formula	http://pubchem.ncbi.nlm.nih.gov [112]

Lead identification approach

Virtual screening

Library of compounds

Structures of receptor

Each compound is sampled in thousands to millions of possible configurations.
The complexes docked are scored on the basis of the binding force. The complex with the highest binding force is selected.

Compounds are docked into the structure of the receptor target by a docking computer program.

Docking results give the protein-ligand binding mode for each compound. The complexes are evaluated and classified according to the scoring function related to binding energy

From the library of compounds, tens of top-scoring predicted ligands (hits) are subsequently tested for activity in an experimental assay.

Molecular docking

Outline of molecular docking process

3D structure of ligand

3D structure of receptor

the ligand is docked into the Binding cavity of the receptor and the putative conformations are explored
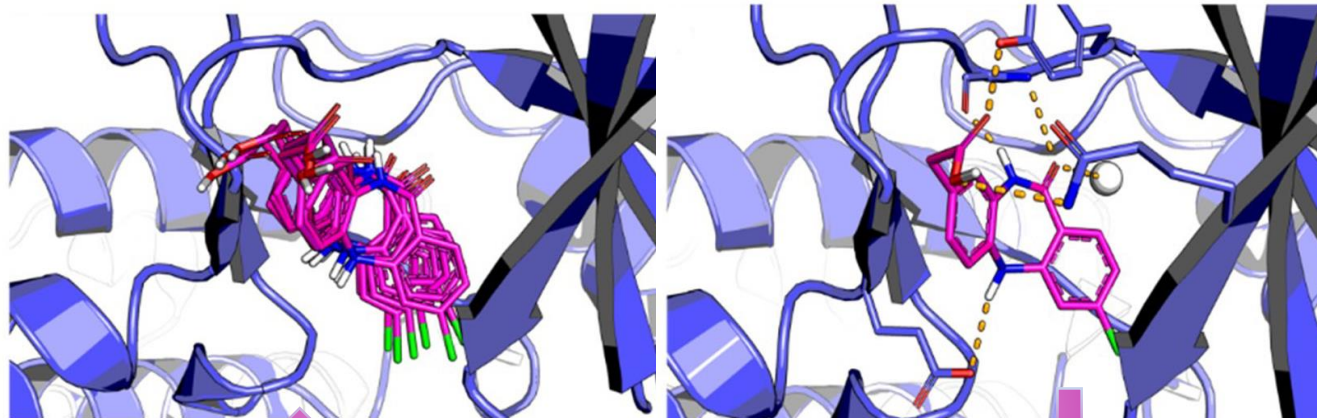
accurate prediction of the interaction energy associated with each of the predicted binding conformations

Docking programs perform these tasks through a cyclical process, in which the ligand conformation is evaluated by scoring function until it converges to the solution of minimum energy

Molecular docking

Conformational research of the ligand
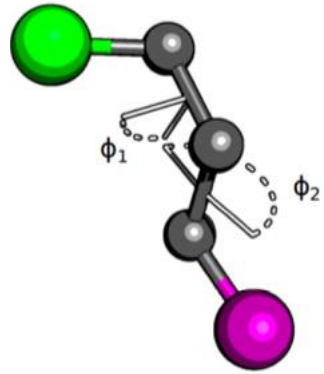
structural parameters of the ligands, such as torsional (dihedral), translational and rotational degrees of freedom, are incrementally modified

systematic search algorithm
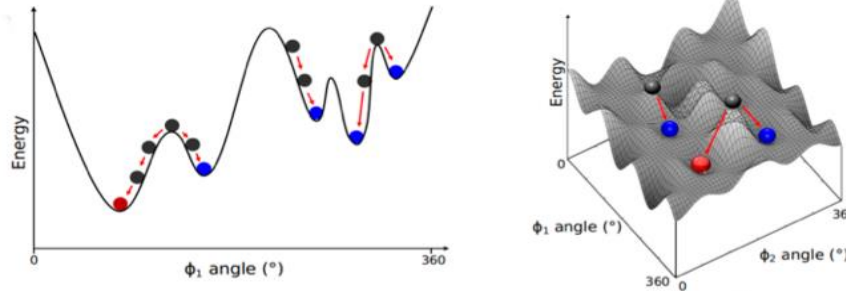
Systematic search methods promote slight variations in the structural parameters of the ligand, gradually changing the conformation of the ligands.

The algorithm explores the energy landscape of the conformational space and, after numerous search and evaluation cycles, converges to the minimum energy solution corresponding to the most likely binding mode. Exploration of the energy landscape is obtained starting from different points of the energy landscape, each of these corresponds to distinct conformations.

Stochastic search algorithm
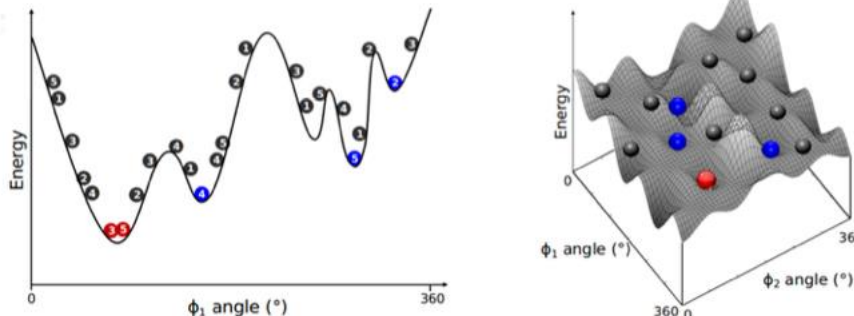
Stochastic methods carry out the conformational search by randomly modifying the structural parameters of the ligands.

The algorithm generates ensembles of molecular conformations and populates a wide range of the energy landscape. This strategy avoids trapping the final solution at a local energy minimum and increases the probability of finding a global minimum

Case study of Cystic fibrosis transmembrane conductance regulator (CFTR) modulators

From gene to CFTR protein structure
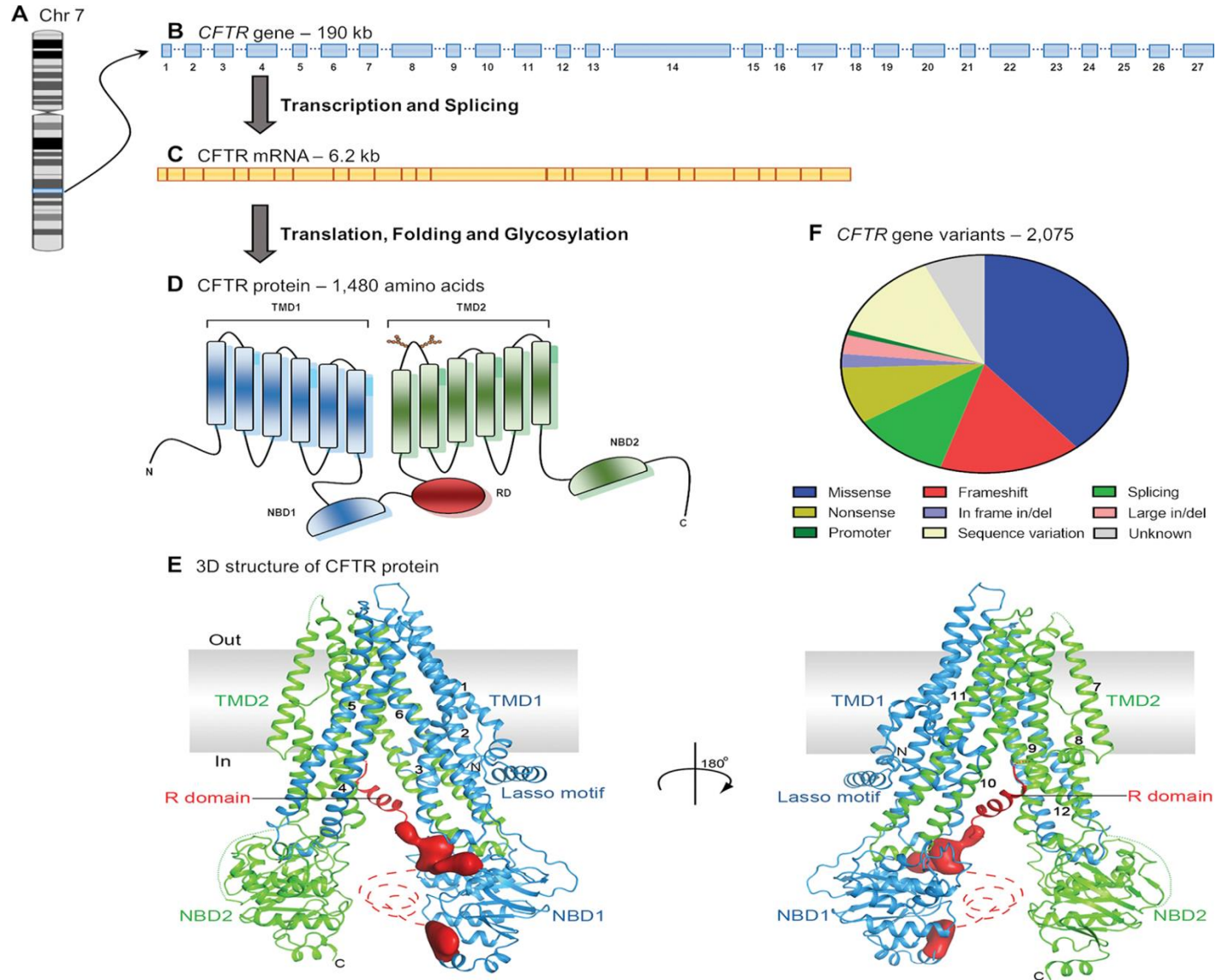
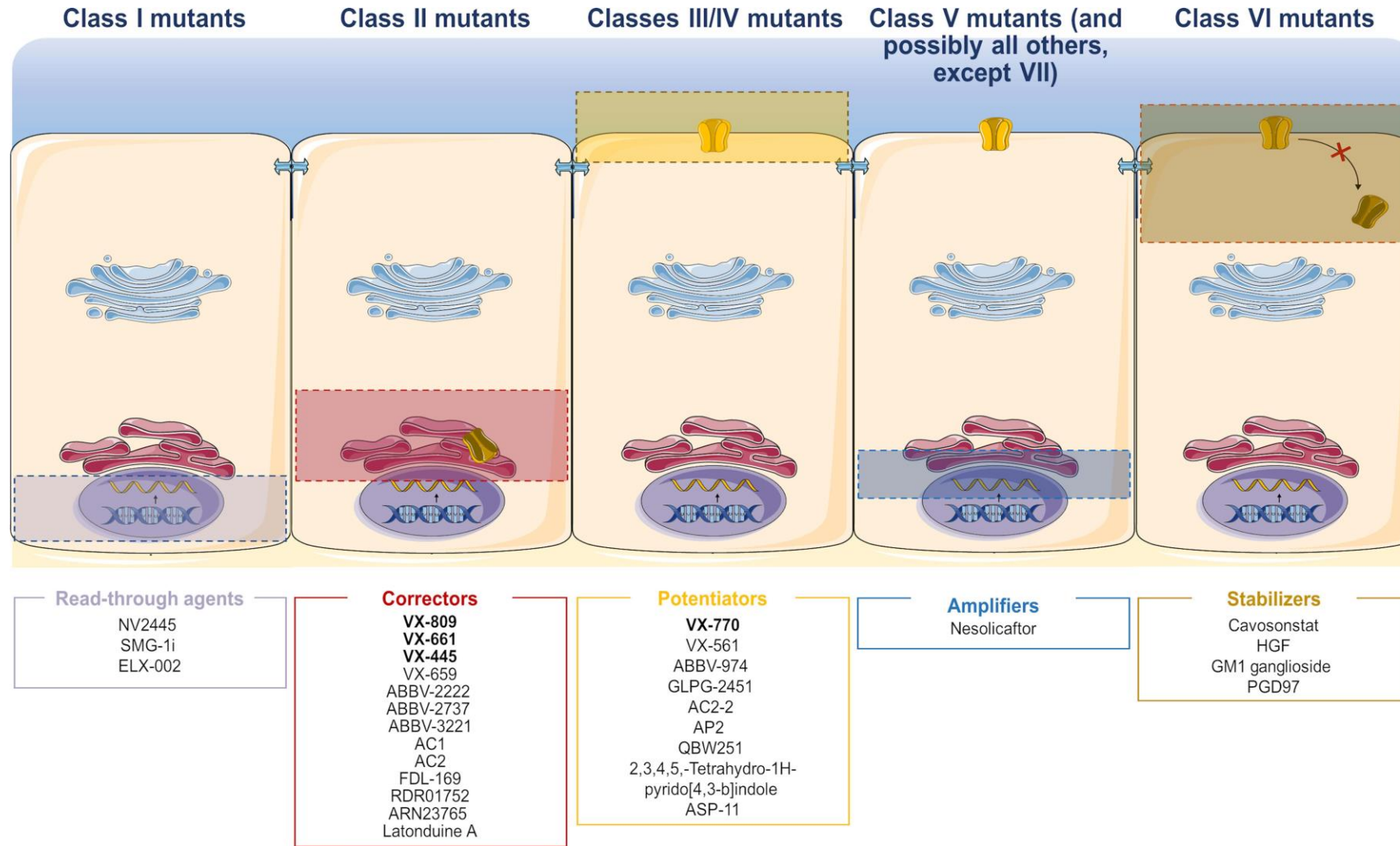# Classes of CF transmembrane conductance regulator (CFTR) mutations

	WT-CFTR	I	II	III	IV	V	VI
Mutation examples		G542X, R553X, R1162X, W1282X	G85E, I507del, F508del, N1303K	S549R, G551D, G1349D	R117H, R334W, D1152H	A455E, 2789+5G>A, 3849+10kbC>T	r.F508del, Q1411X
Defect types		Protein synthesis	Protein traffic	Channel gating	Channel conductance	Protein abundance	PM protein stability
		Minimal function			Residual function		
Clinical features (global aspects)		More severe phenotype			Milder phenotype		

I   no production of full-length protein
II defective folding and trafficking
III defective gating

IV reduced anion conductance
V   reduced protein production
VI reduced stability at the PM

High-throughput screening for modulators of CFTR activity

- Numerous assays and high-throughput screens (HTS) have been developed and optimized to screen drug-like compound libraries and identify CFTR modulators.

- These specialized small molecules target the root cause of CF to enhance or even restore the expression, function, and stability of a defective CFTR in distinct manners, and they have been classified into five main groups depending on their effects on CFTR mutations.

Site of action of the different CFTR modulator drugs

Class I mutants — Read-through agents: NV2445, SMG-1i, ELX-002

Class II mutants — Correctors: **VX-809**, **VX-661**, **VX-445**, VX-659, ABBV-2222, ABBV-2737, ABBV-3221, AC1, AC2, FDL-169, RDR01752, ARN23765, Latonduine A

Classes III/IV mutants — Potentiators: **VX-770**, VX-561, ABBV-974, GLPG-2451, AC2-2, AP2, QBW251, 2,3,4,5,-Tetrahydro-1H-pyrido[4,3-b]indole, ASP-11

Class V mutants (and possibly all others, except VII) — Amplifiers: Nesolicaftor

Class VI mutants — Stabilizers: Cavosonstat, HGF, GM1 ganglioside, PGD97

CFTR modulator drugs may be grouped according to their actions on CFTR mutations

Read-through agents
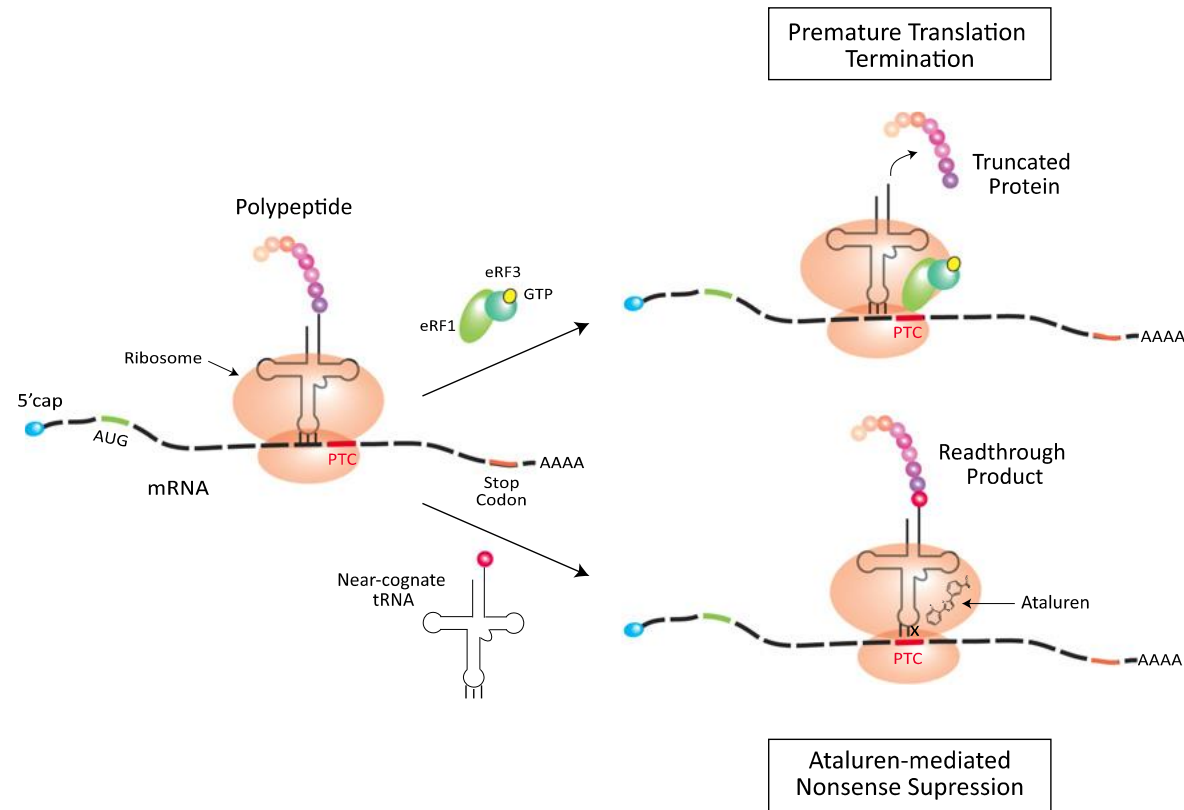
Class I mutations when a ribosome encounters a premature stop codon (PTC) due to a nonsense mutation, eukaryotic release factors in complex with GTP are recruited and translation is terminated prematurely, generating a truncated protein.
Ataluren interacts with the ribosome and facilitates the recruitment of near-cognate tRNAs, which suppresses the non-sense mutation and allows for the readthrough of a PTC and synthesis of a full-length protein

Correctors
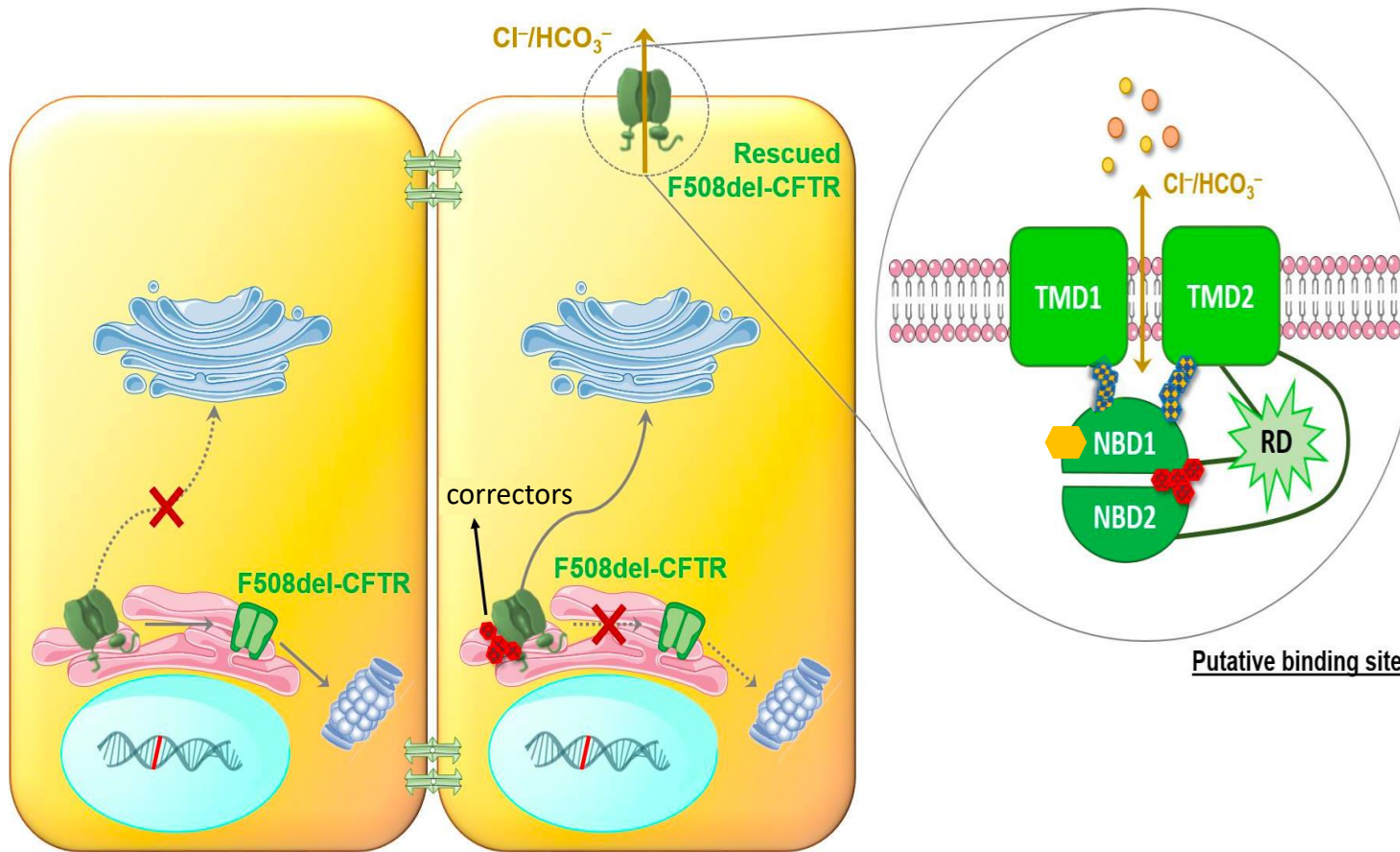
Correctors are classified based on their molecular targets in the CFTR structure

class 1 correctors stabilize NBD1-TMD1 and/or NBD1-TMD2

class 2 correctors stabilize NBD2 and its interfaces with other CFTR domains

class 3 correctors directly stabilize NBD1

This classification may be useful to evaluate corrector combinations with complementary mechanisms of action

Correctors are small molecules that rescue CFTR mutants with a traffic defect to the PM. These compounds may act by distinct mechanisms, they usually enhance protein conformational stability during the ER folding process. The structure of wild type (WT) CFTR shows that F508 resides on the surface of NBD1, where it makes extensive interactions with the cytosolic region of TM helix 11 and intracellular loop 4. These interactions are critical for both CFTR folding and coupling of ATP dependent NBD dimerization to pore opening, suggesting that disruption of these interactions may underlie both trafficking and functional defects in Δ508 CFTR.

Regulation of CFTR gating
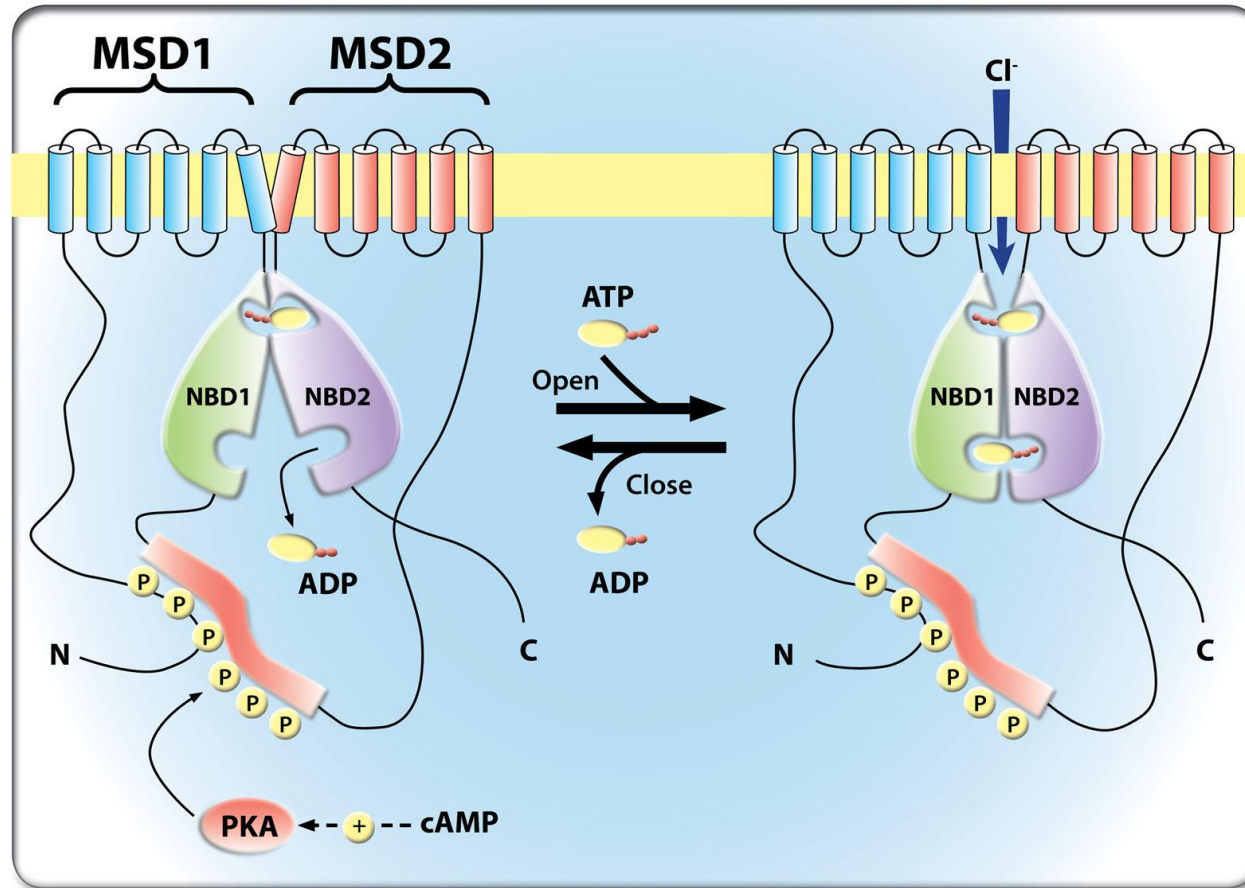
CFTR gating is regulated by the R domain which contains multiple PKA and PKC phosphorylation sites and which physically interacts with NBD1. CFTR uses the energy of ATP binding and hydrolysis to drive ligand-induced conformational changes in the protein that lead to the regulated opening and closing (gating) of the channel 'pore'.

Phosphorylation in CFTR channel activity regulation
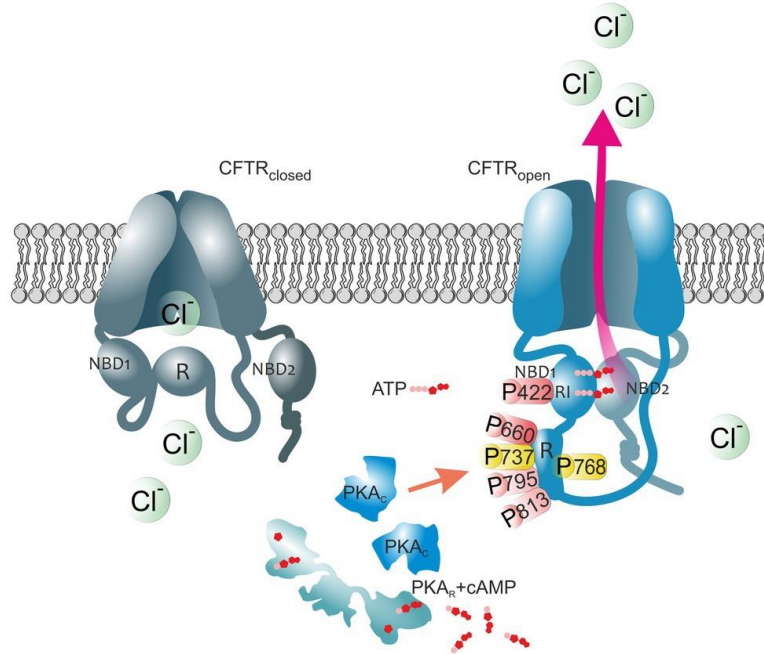
In the closed CFTR state, the NBD1 interacts with the R domain creating steric hindrances that prevent it from dimerization with NBD2. Phosphorylation causes large conformational changes that remove the R region from its inhibitory position and allow NBDs dimerization to occur.
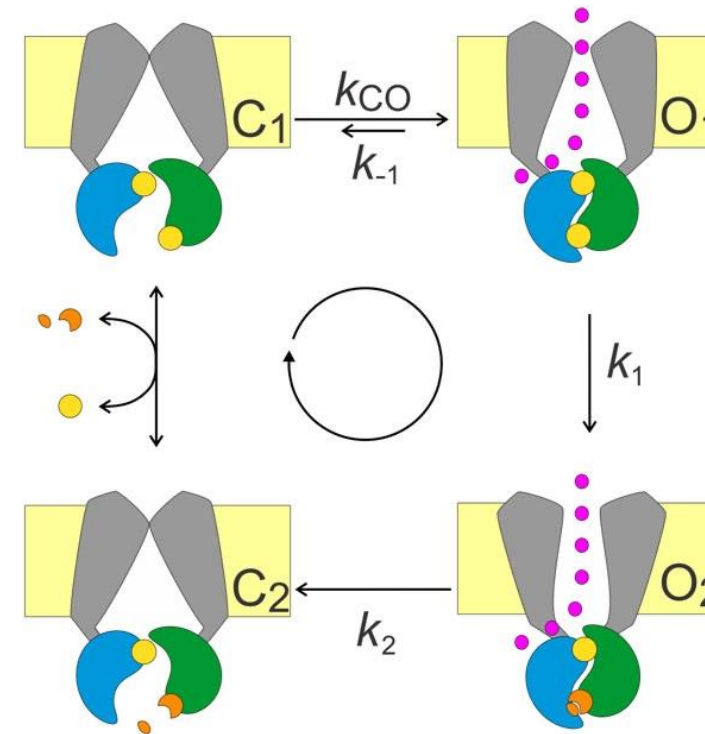
The channel enters a stable open state (O1) following the formation of a head-to-tail NBD dimer, with ATP bound at the two ATP binding sites formed at the interface.
Upon hydrolysis of the ATP bound at site 2, the channel closes via a short-lived post-hydrolytic (O2). Mutations disrupt any critical steps in this gating process (opening and closing) and may lead to CFTR dysfunction.

Potentiators
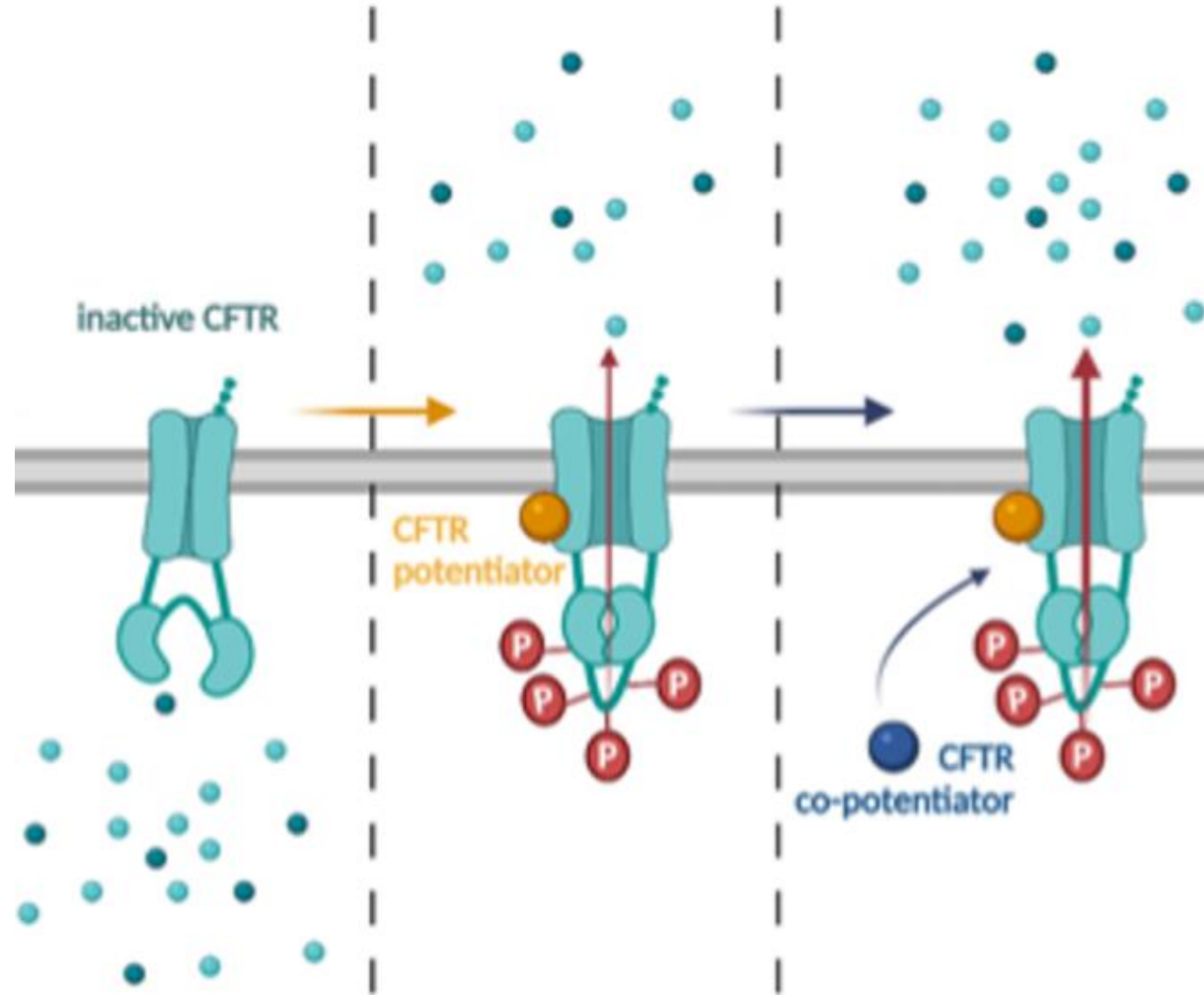
CFTR potentiators interact directly with CFTR to promote function, in a phosphorylation-dependent, but ATP-independent manner. Co-potentiators also interact with CFTR, to further stimulate function but only in the presence of CFTR potentiators.

Amplifier

The amplifier stabilizes the mRNA, which results in an increase in available CFTR mRNA transcripts.
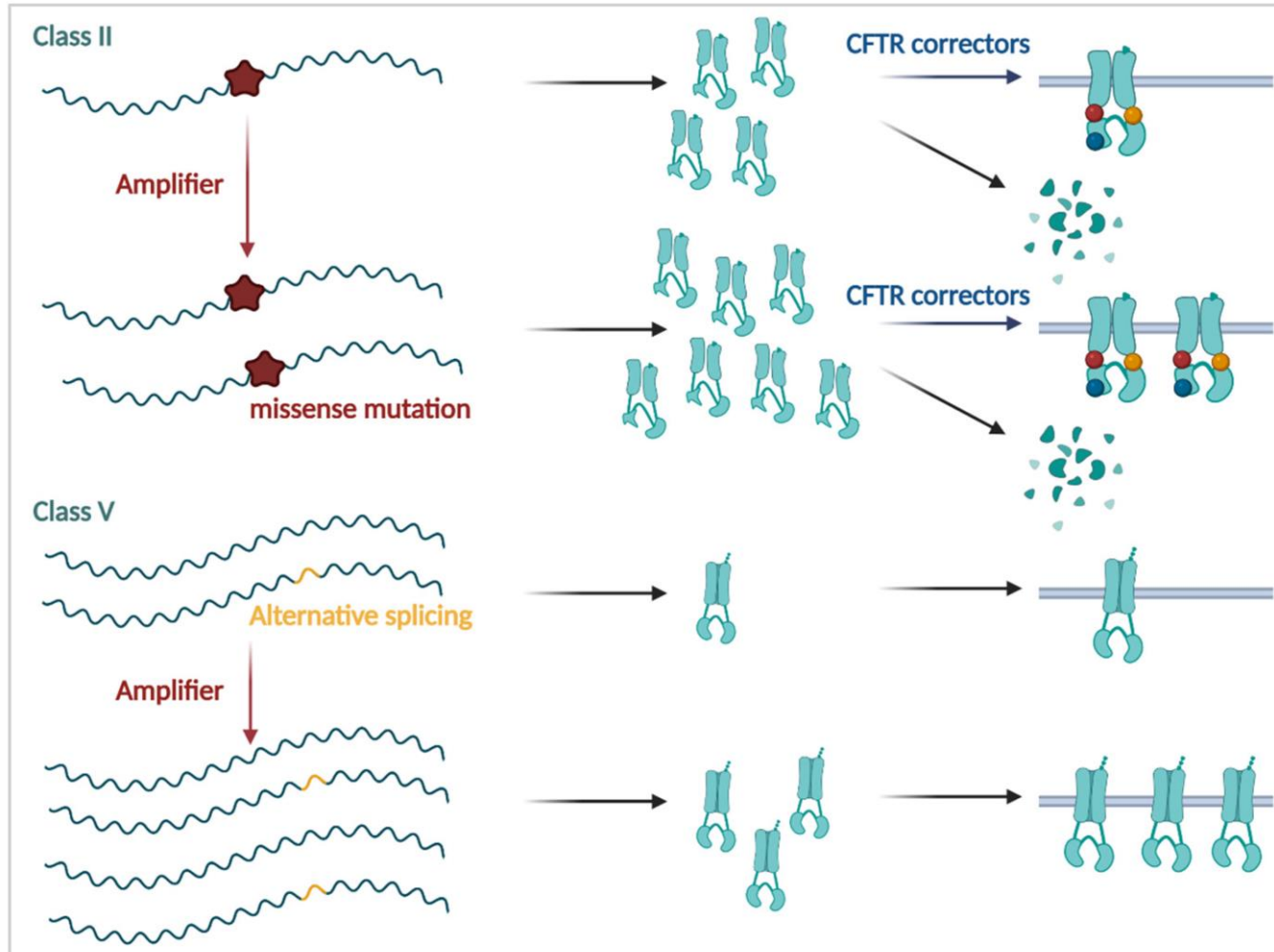
Class II mutations are misfolded and only a fraction reaches the plasma membrane, even in the presence of CFTR correctors. Amplifiers provide more immature protein that can be rescued with CFTR modulators.

Class V introduce cryptic splice sites, resulting in a mix of normal and alternatively spliced CFTR mRNA. Enhancing the number of correct mRNA transcripts results in more CFTR protein at the plasma membrane.

Mechanisms of action of CFTR modulators approved for people with CF
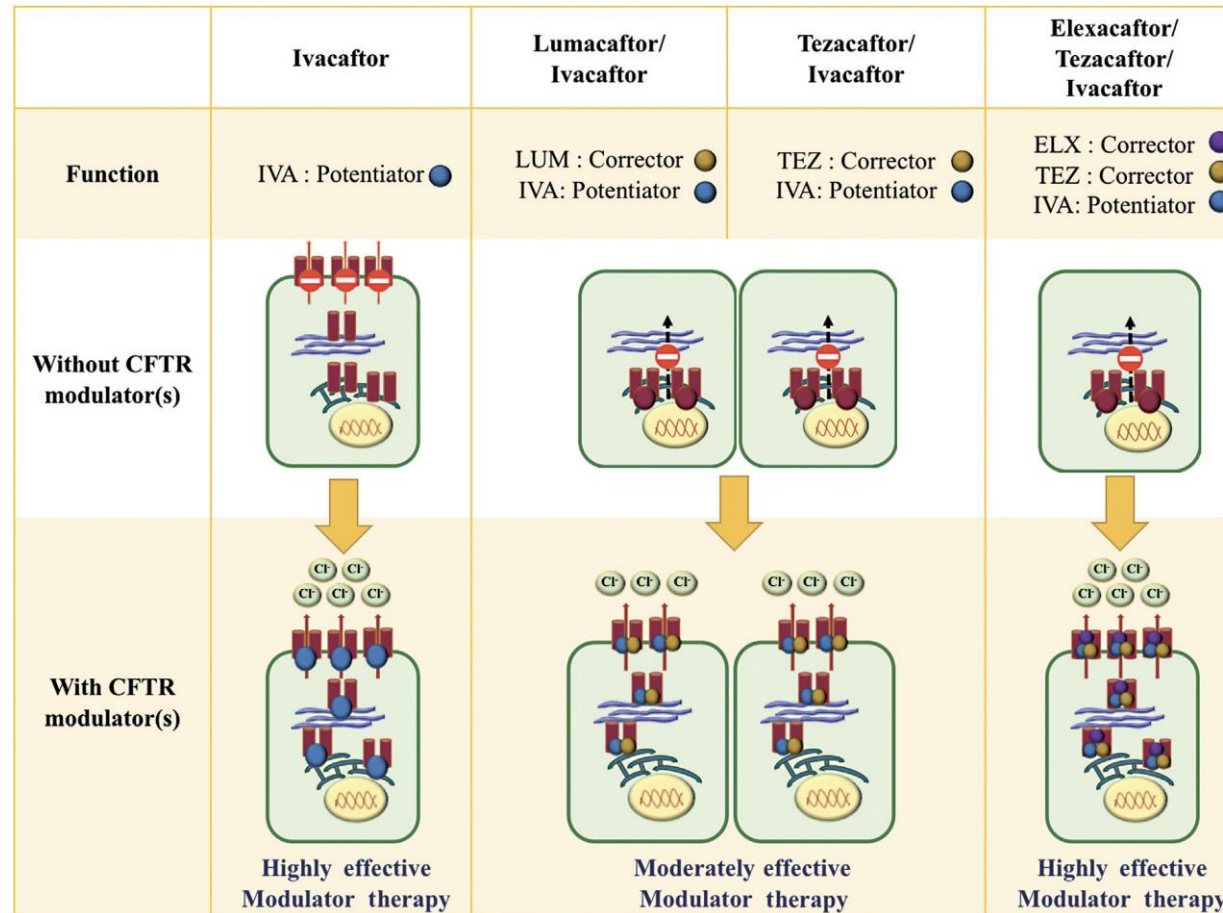
IVA increases CFTR opening frequency and ion conductance.

LUM, TEZ, and ELEX are correctors that have an additive effect to facilitate the intracellular maturation and trafficking of the CFTR protein to increase the amount of CFTR proteins brought to the cell surface.

 LUM/IVA and TEZ/IVA are moderately effective in F508del homozygous patients, but TEZ/IVA is highly effective in patients with one F508del mutation.

A triple combination of ELX/TEZ/IVA is considered highly effective: high level of CFTR function restoration (major decrease in sweat chloride concentration) and rapid and major clinical improvement

~9 Å

NBD1
disordered

lumacaftor

~9 Å
disordered

elexacaftor

3.7 Å
cracked

lumacaftor
elexacaftor

2.8 Å
corrected

Trikafta

3.0 Å
corrected

Δ508 + elexacaftor

Δ508 + elexacaftor + lumacaftor

Δ508 + elexacaftor + tezacaftor + ivacaftor

Fiedorczuk, Karol, and Jue Chen. "Molecular structures reveal synergistic rescue of Δ508 CFTR by Trikaftamodulators."
Science vol. 378,6617 (2022): 284-290.

Projects funded by Fondazione Fibrosi Cistica

FFC#10/2019
Rescuing defective CFTR by applying a drug repositioning strategy based on computational studies, surface plasmon resonance and cell-based assays.

FFC#11/2018
Rescuing defective CFTR by applying a drug repositioning strategy based on computational studies, surface plasmon resonance and cell-based assays.

FFC#6/2014
Development of novel methodologies for the identification of cftr-targeted drugs: a multidisciplinary approach using real-time surface plasmon resonance interaction assay
supported by bioinformatic strategies on HPC infrastructures

PI: Prof. Marco Rusnati

Centres Involved: University of Brescia, University of Genova and CNR

Projects funded by the Cystic Fibrosis Foundation

Novel methodologies for the identification of cftr-targeted drugs were developed by a multidisciplinary approach using real-time surface plasmon resonance interaction assay supported by bioinformatic strategies on HPC infrastructures
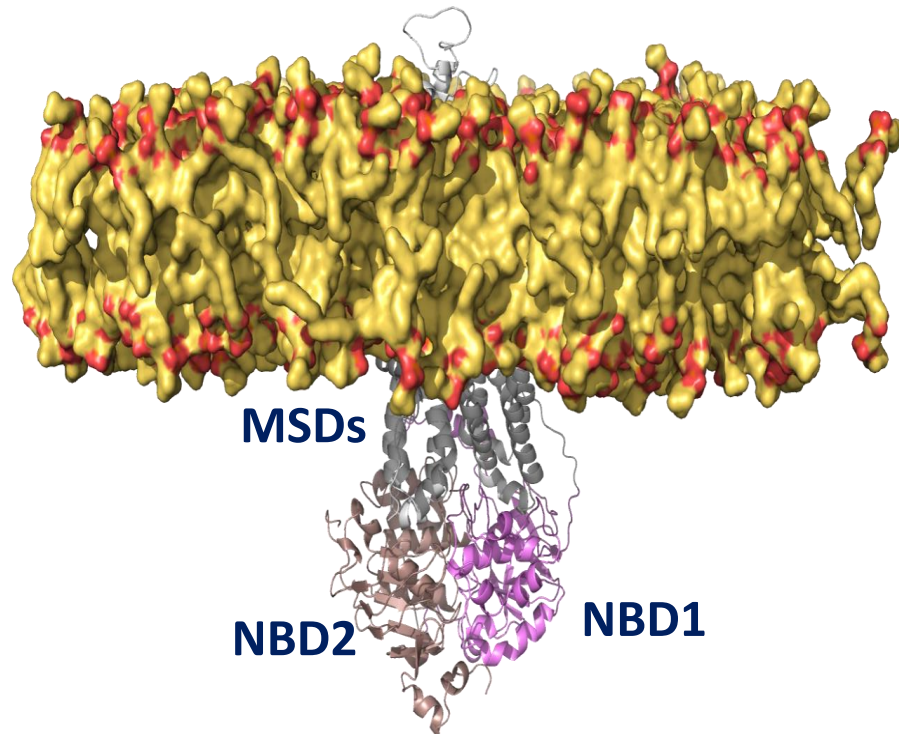
A computational model of F508del CFTR obtained by homology modelling and molecular dynamics. The F508del CFTR is embedded in the lipid environment
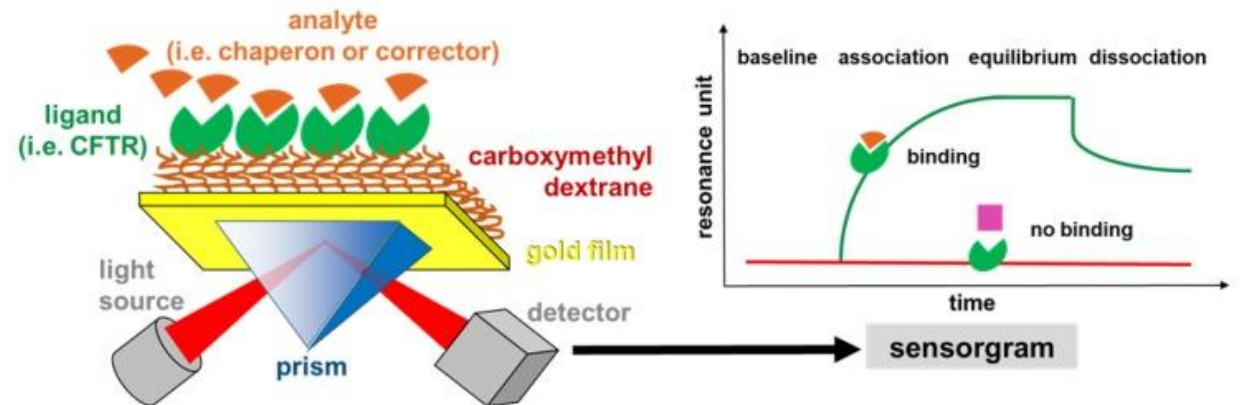
The biosensor includes F508del-CFTR in a lipid environment and has a wider analytical potential than the previous one based on CFTR

The computational model was used to explore the binding modes of some known F508del-CFTR-rescuing drugs and of a new aminoarylthiazole
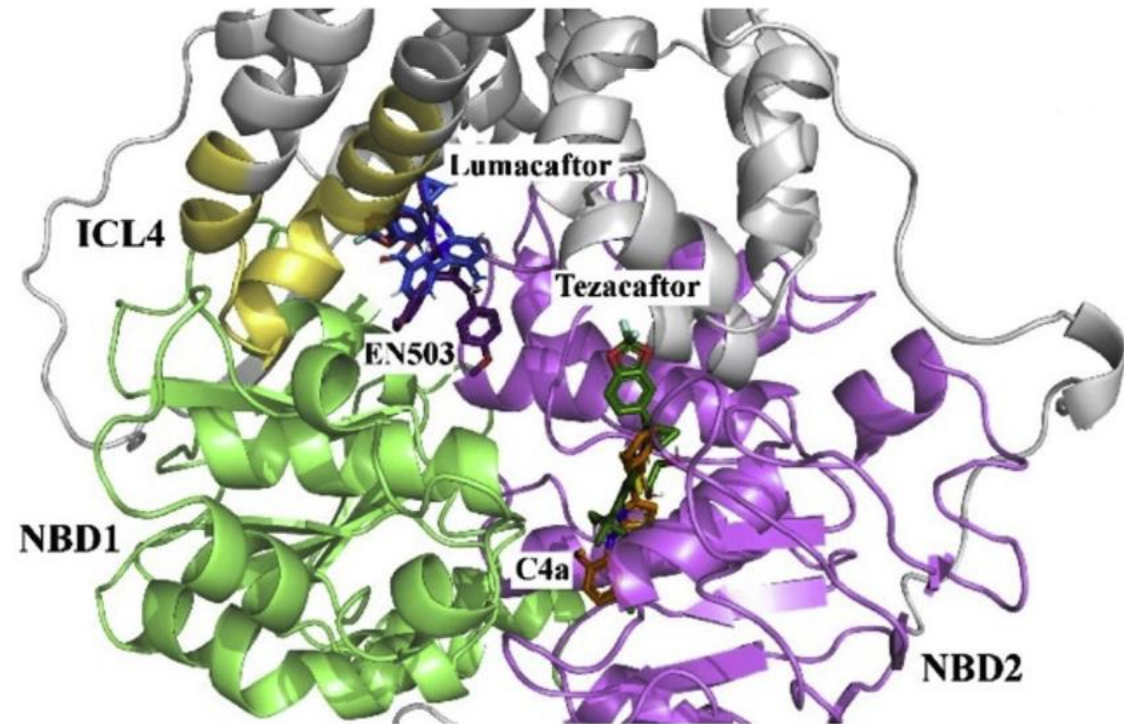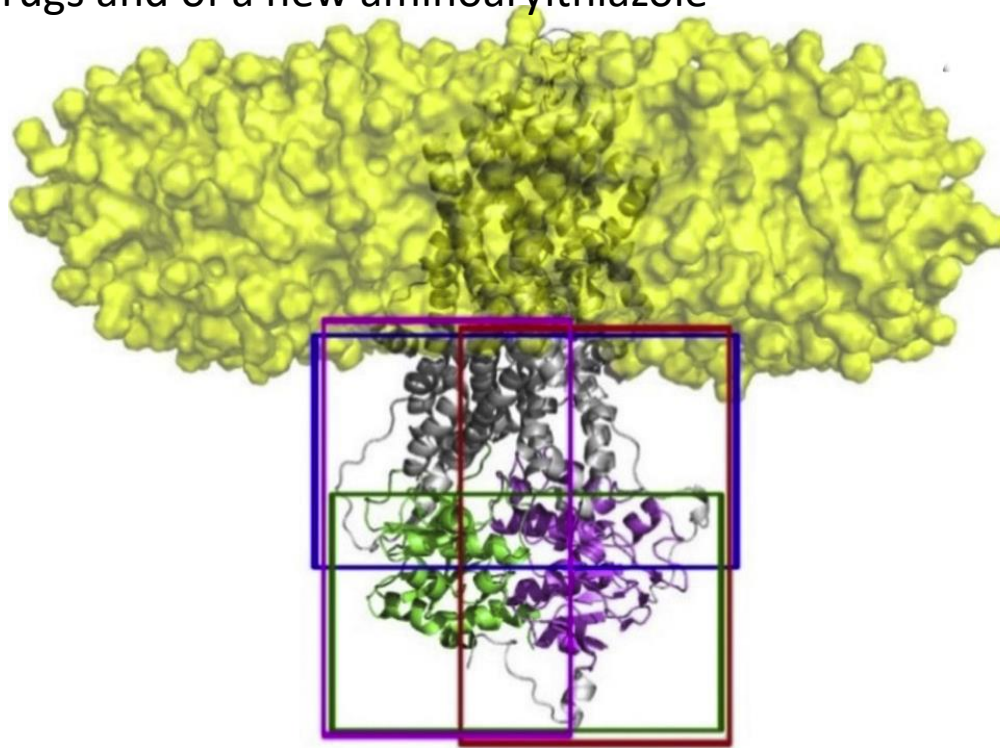
The multiple boxes method can cover the entire protein to explore the binding site of ligand. Red box: NBD1-MSDs region partially overlapping NBD2-MSDs; purple box: NBD2-MSDs region partially overlapping NBD1-MSDs; green box: NBD1-NBD2 region; blue box: MSDs, NBD1 and NBD2 regions.

Docking Poses of compounds into apo F508del-CFTR. Lumacaftor (blue), Tezacaftor (dark green), C4a (orange) and EN503A (purple) are shown as stick, NBD1 (green), NBD2 (violet), ICL4 (gold) as cartoon

Molecular dynamics (MD) is a computational technique that simulates the dynamic behaviour of molecular systems as a function of time. The integration of Newton's equations calculates the movements of atoms over time:

Fi(t) force exerted on atom i at time t

ri(t) vector position of the atom i at time t $\frac{d^2 r_i(t)|}{dt^2} = \frac{F_i(t)}{m_i}$
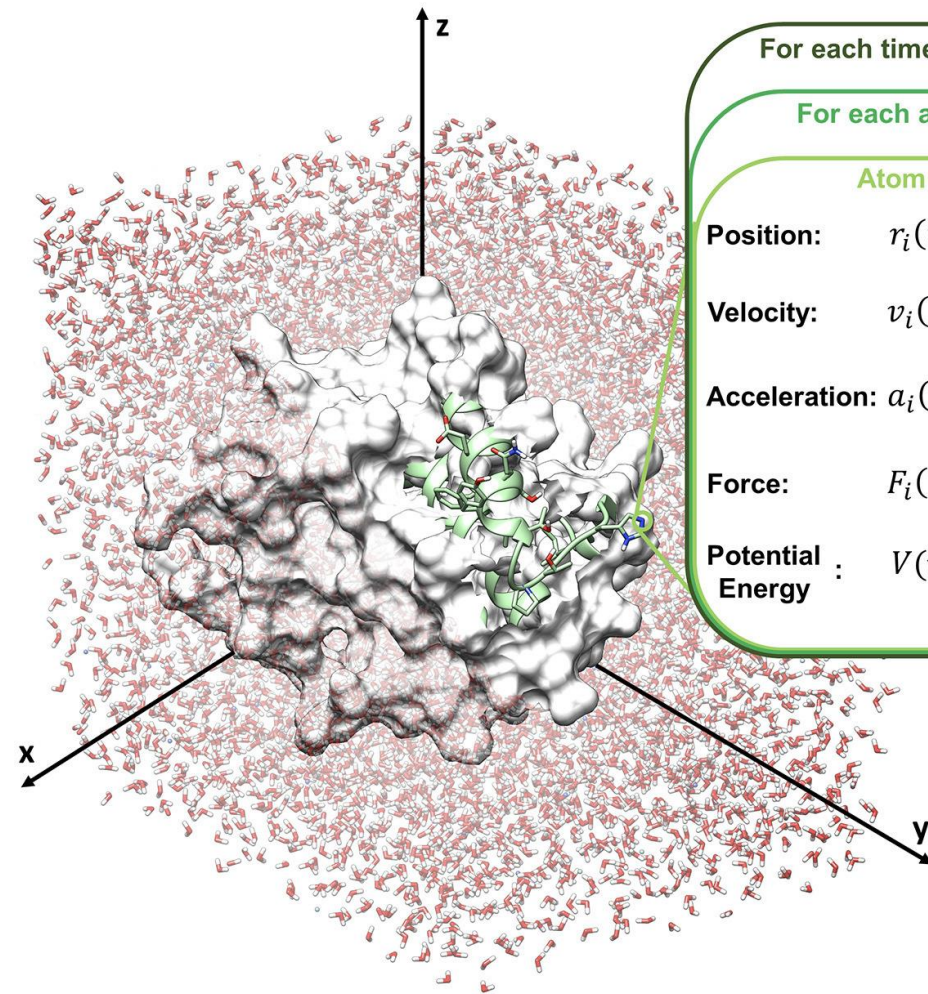
mi mass of the atom

For each timestep (δt):

For each atom:

Atom *i* at time *t*:

Position: $r_i(t) = (x_i(t), y_i(t), z_i(t))$

Velocity: $v_i(t) = \frac{dr_i(t)}{dt}$

Acceleration: $a_i(t) = \frac{d^2 r_i(t)}{dt^2} = \frac{F_i(t)}{m_i}$

Force: $F_i(t) = -\frac{dV(r(t))}{dr_i(t)}$

Potential Energy : $V(r(t))$

Exploring the stability of ligand binding modes to CFTR by molecular dynamics simulation

Time is partitioned into time steps (δt), which are used to propagate the system forward in time. Several integration algorithms are available, which derive Newton's equations by a discrete-time numerical approximation. An example to compute the position and velocity of an atom i at the time step t+δt, starting from step t:

$$r_i(t + \delta t) = r_i(t) + v_i(t)\,\delta t + \frac{1}{2} a_i(t)\,\delta t^2$$

$$v_i(t + \delta t) = v_i(t) + \frac{1}{2}\left[a_i(t) + a_i(t + \delta t)\right]\delta t$$

ri(t), vi(t) and ai(t) are the position, velocity and acceleration of atom i at time t

ri(t+δt), vi(t+δt) and ai(t+δt) are the position, velocity and acceleration of atom i at time t+δt.

For each timestep (δt):

For each atom:

Atom *i* at time *t*:

Position: $r_i(t) = (x_i(t), y_i(t), z_i(t))$

Velocity: $v_i(t) = \dfrac{\mathrm{d}r_i(t)}{\mathrm{d}t}$

Acceleration: $a_i(t) = \dfrac{\mathrm{d}^2 r_i(t)}{\mathrm{d}t^2} = \dfrac{F_i(t)}{m_i}$

Force: $F_i(t) = -\dfrac{\mathrm{d}V(r(t))}{\mathrm{d}r_i(t)}$

Potential Energy : $V(r(t))$
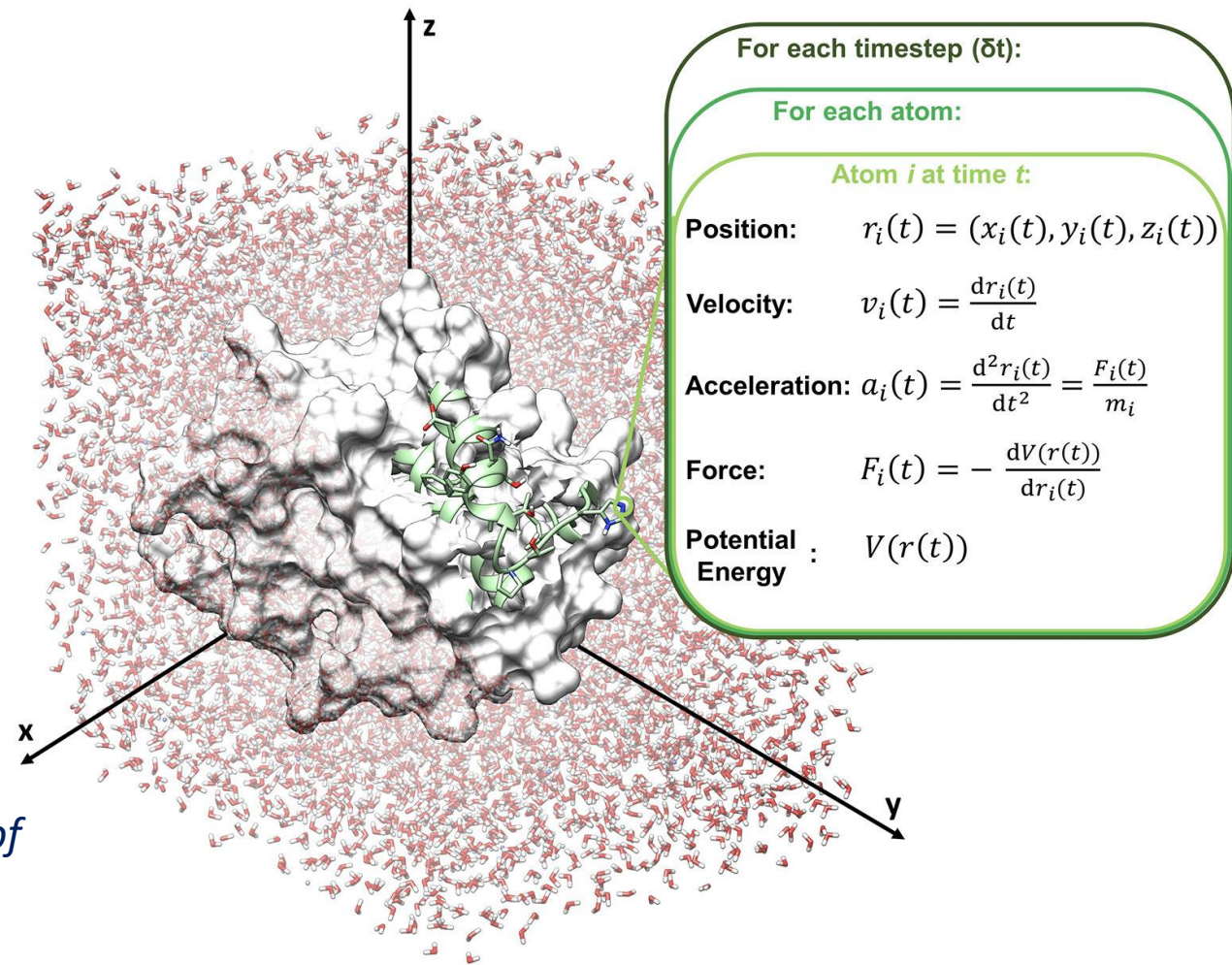
Exploring the stability of ligand binding modes to CFTR by molecular dynamics simulation

Acceleration is calculated from the forces acting on atom i according to Newton's second law, and forces are computed from the force field, according to the following equation:

$$a_i(t) = \frac{d^2 r_i(t)}{dt^2} = \frac{F_i(t)}{m_i} = -\frac{dV(r(t))}{m_i dr_i(t)}$$

V(r(t)) is the potential energy function retrieved by the force field
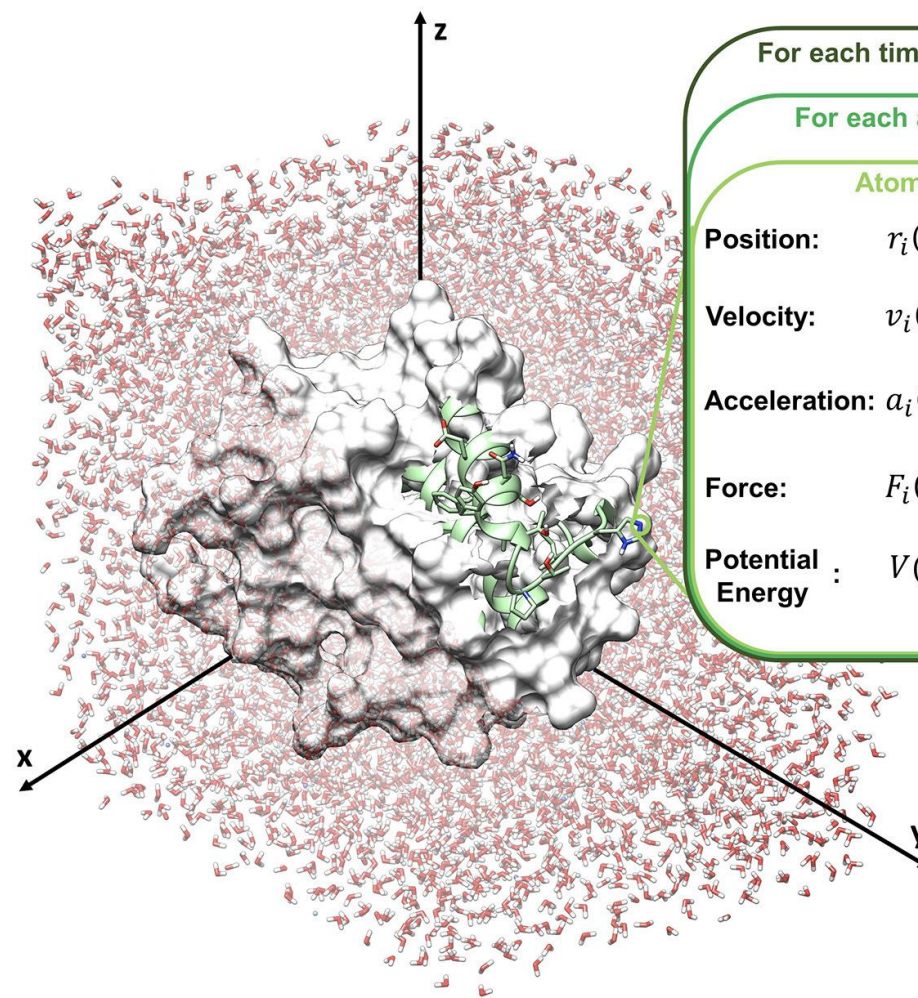
For each timestep (δt):

For each atom:

Atom *i* at time *t*:

Position: $r_i(t) = (x_i(t), y_i(t), z_i(t))$

Velocity: $v_i(t) = \frac{dr_i(t)}{dt}$

Acceleration: $a_i(t) = \frac{d^2 r_i(t)}{dt^2} = \frac{F_i(t)}{m_i}$

Force: $F_i(t) = -\frac{dV(r(t))}{dr_i(t)}$

Potential Energy : $V(r(t))$

Exploring the stability of ligand binding modes to CFTR by molecular dynamics simulation

The potential energy function used to derive atomic forces for molecular movement is the sum of bonded (intramolecular) and non-bonded energy terms.

$$V(r(t)) = \sum_{bonds} k_r (r - r_{eq})^2 \qquad bond$$

$$+ \sum_{angles} k_\theta (\theta - \theta_{eq})^2 \qquad angle$$

$$+ \sum_{dihedrals} k_\phi (1 + \cos[n\phi - \gamma]) \qquad dihedral$$

$$+ \sum_{impropers} k_\omega (\omega - \omega_{eq})^2 \qquad improper$$

$$+ \sum_{i<j}^{atoms} \varepsilon_{ij} \left[\left(\frac{r_m}{r_{ij}} \right)^{12} - 2 \left(\frac{r_m}{r_{ij}} \right)^6 \right] \qquad van\ der\ Waals$$
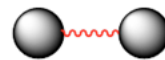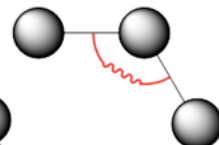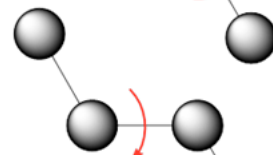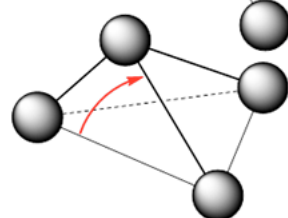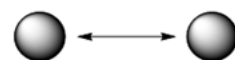
$$+ \sum_{i<j}^{atoms} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \qquad electrostatic$$
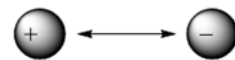
r: bond length;

θ : atomic angle;

φ : dihedral angle;

ω: improper dihedral angle;

rij : the distance in between atom i and j;

kr, kθ, kφ, and kω : force constants;

req, θeq and ωeq are equilibrium positions;

dihedral term is a periodic term characterized by a force constant (kω), multiplicity (n), and phase shift (γ);

εij is related to the Lennard–Jones well depth;

rm is the distance at which the potential reaches its minimum;

qi and qj are the charges on the respective atoms;

ε0 is the dielectric constant.

Exploring the stability of ligand binding modes to CFTR by molecular dynamics simulation

Forces acting on individual atoms are then used for the calculation of accelerations and velocities by Newton's law of motion and atom positions are updated after each time step

RMSD function quantifies the difference between two poses of the same molecule
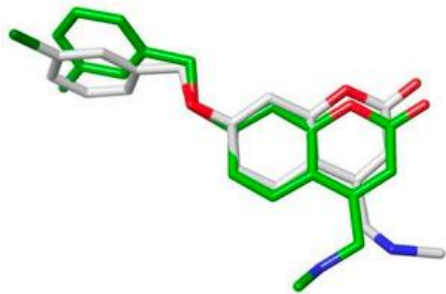
$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2}$$

δ_i is the distance between atom i and either a reference structure
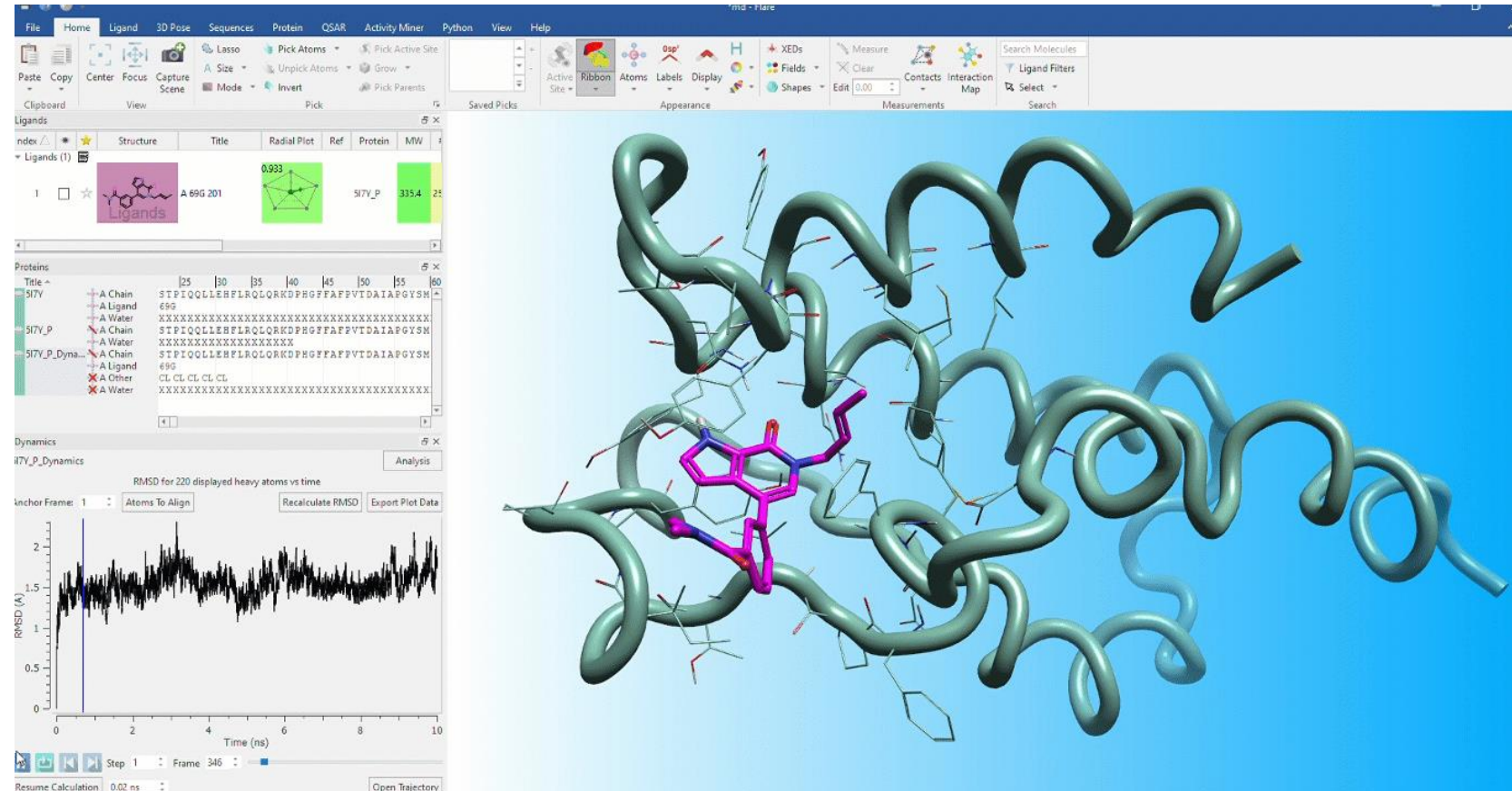
Exploring the stability of ligand binding modes to CFTR by molecular dynamics simulation

RMSD of the F508del-CFTR/Lumacaftor complex shows that it reaches stability at 12 ns
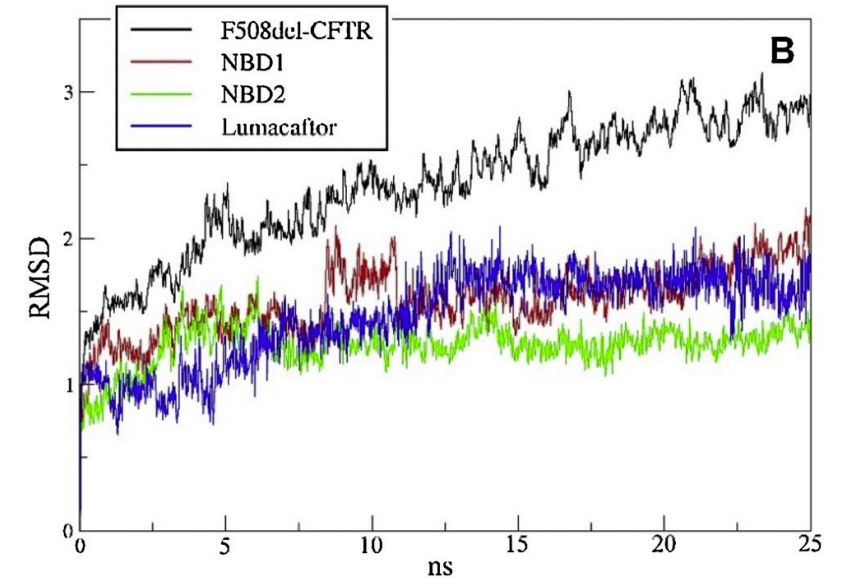
Lumacaftor interacts with an S492, W496, R560 of NBD1 and with W1063 of ICL4. These predictions are supported by recent experimental data suggesting a binding pocket for Lumacaftor between NBD1 and ICL4
(R.P. Hudson and et al. Direct binding of the corrector VX-809 to human CFTR NBD1: evidence of an allosteric coupling between the binding site and the NBD1:CL4 interface, Mol. Pharmacol. 92 (2) (2017) 124–135)
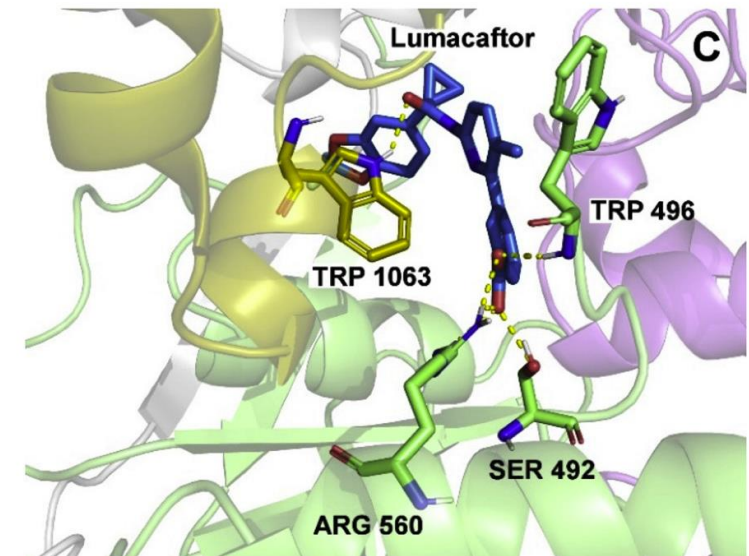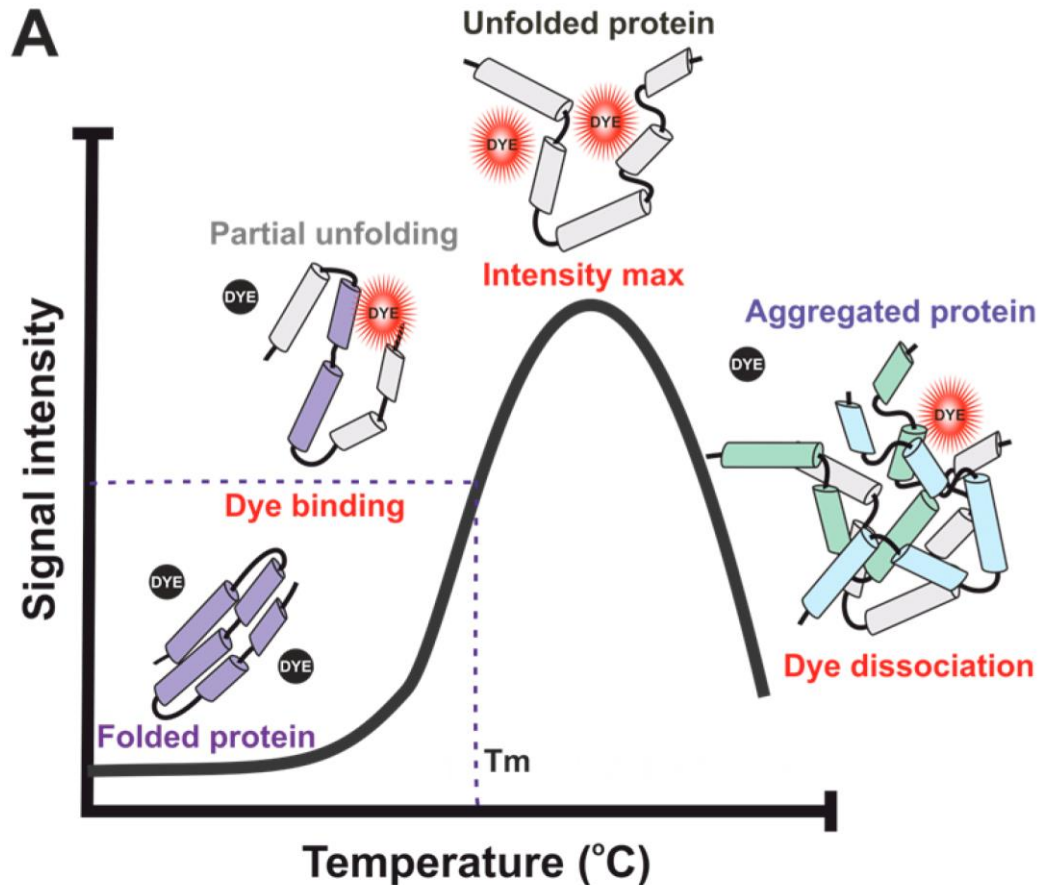
Typical thermal stability assay (TSA) data using external fluorescent dye. (**A**) Increase in temperature causes protein unfolding, enabling dye binding and increase in fluorescence. Melting temperature (T_m) is observed as a half-maximum signal, and further increase in temperature leads to dye dissociation due to protein aggregation.
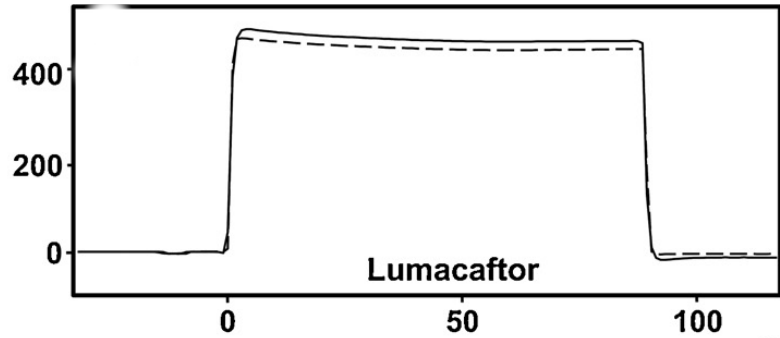
Sensorgram showing the binding of Lumacaftor

The effect on the stability of F508CFTRdel was evaluated by increasing the concentration of compounds: blue circles: 10 µM; red squares 40 µM; green triangles: 80 µM; purple inverted triangles: 120 µM; orange diamonds: 180 µM.

The activity of compounds as correctors. The bar graphs report F508del-CFTR activity in CFBE41o-cells after treatment with the indicated compounds for 24 h at 0.08, 0.4, 2.0 and 10 µM. The dashed line indicates the level of activity in cells treated with a vehicle alone

Rescuing defective CFTR by applying a drug repositioning strategy based on computational studies, surface plasmon resonance and cell-based assays.

Drug repositioning

Drug repositioning is a strategy to accelerate the drug development process by seeking new indications for already approved drugs, rather than discovering de novo active compounds.
The drug repositioning method reduces the cost and the time required for drug development, as information regarding the safety, pharmacokinetics, and formulation of existing drugs is already available.
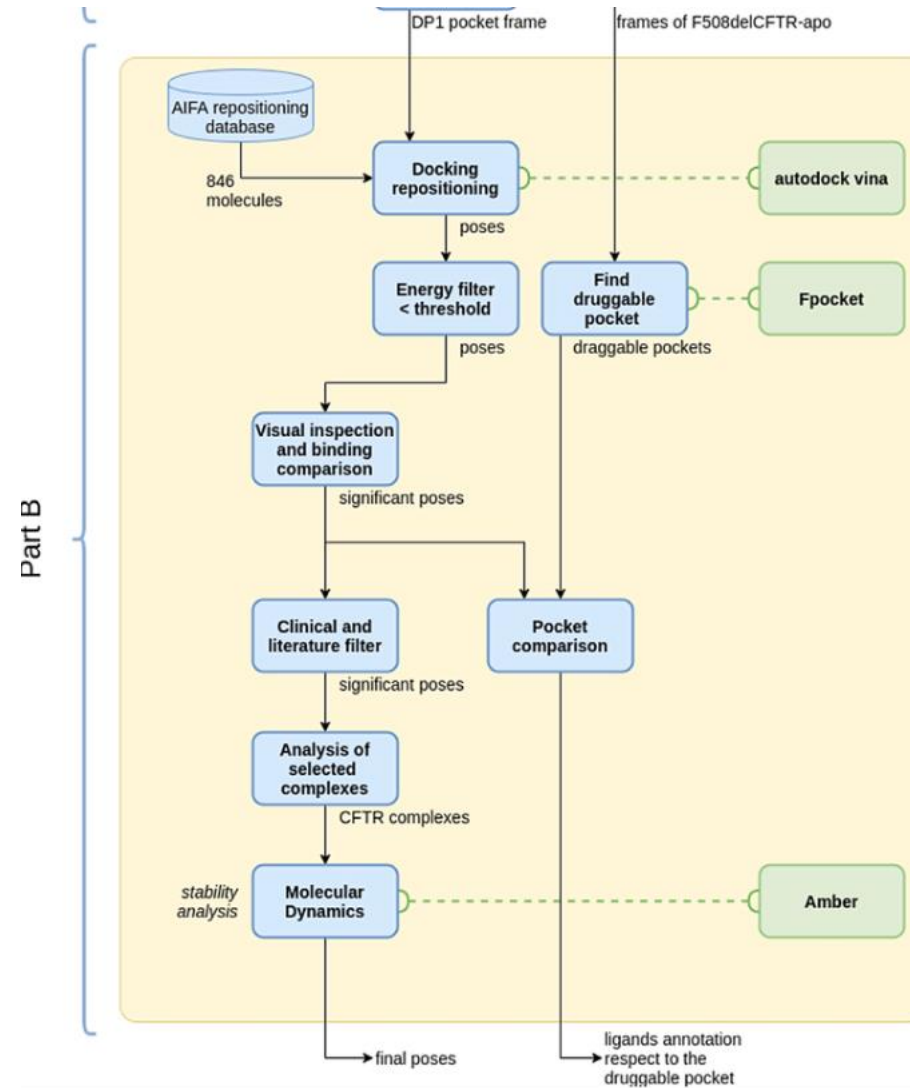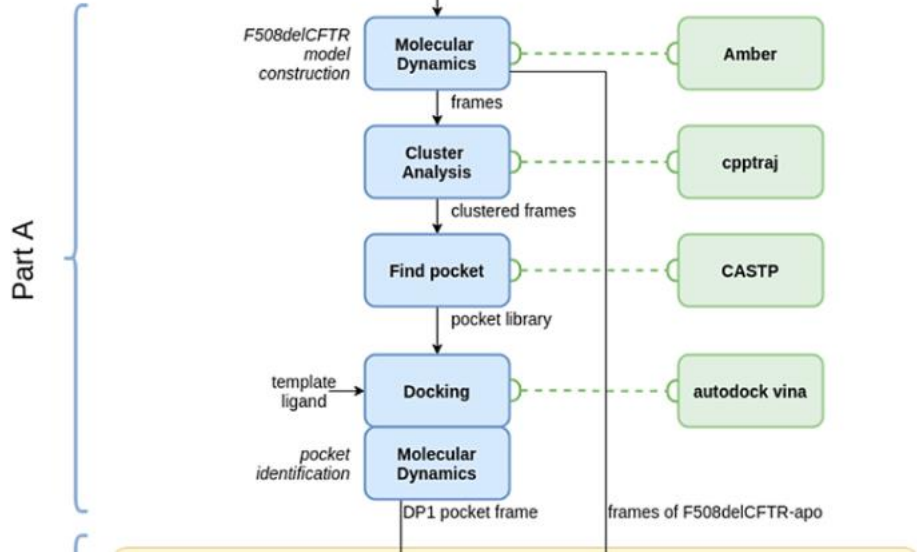
The druggable Pocket1 described before was used in the drug's repositioning pipeline to screen 846 compounds of the AIFA database. Docking repositioning results were firstly filtered by energy, taking into account only those poses with docking energies like that of Lumafactor.

ManXDR: a manually curated ligand database for drug repositioning
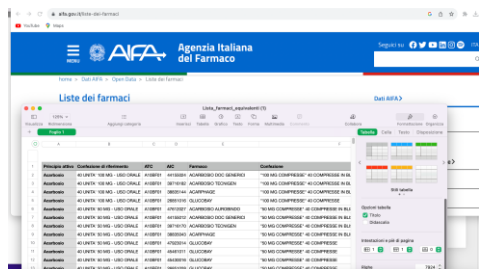
smiles format of the molecule

Simplified **M**olecular **I**nput **L**ine **E**ntry **S**ystem

SMILES:

CC(C)C1=CC=C(C=C1)C(C)C(=O)O
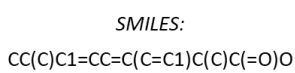
3D modeling
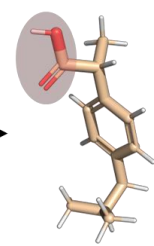
Charging the structure at physiological pH 7,4
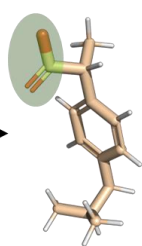
An in-house ligand dataset was manually curated. The molecules were retrieved from the Agenzia Italiana del FArmaco – AIFA and filtered excluding the not appropriate ones, such as hormones, inorganic salts, contrast agents, peptides, and proteins.

ManXDR database

846 molecules
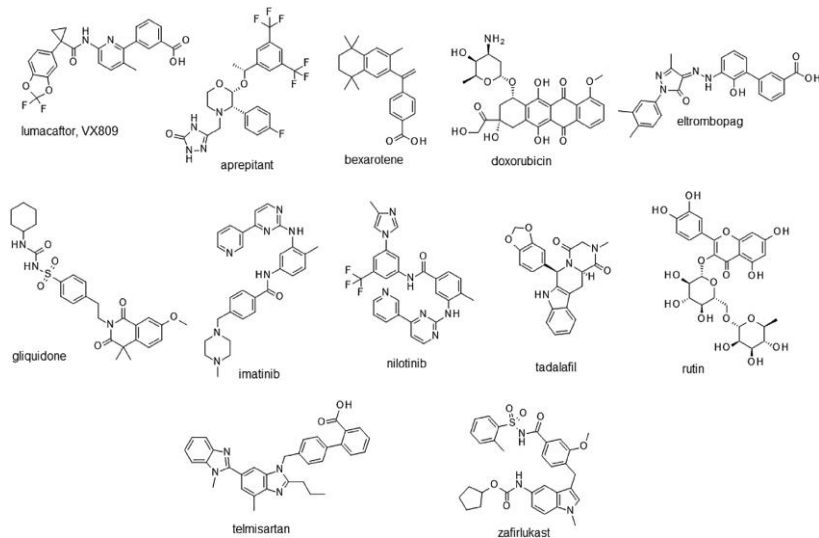
Chemical structures of the 11 drug repositioned
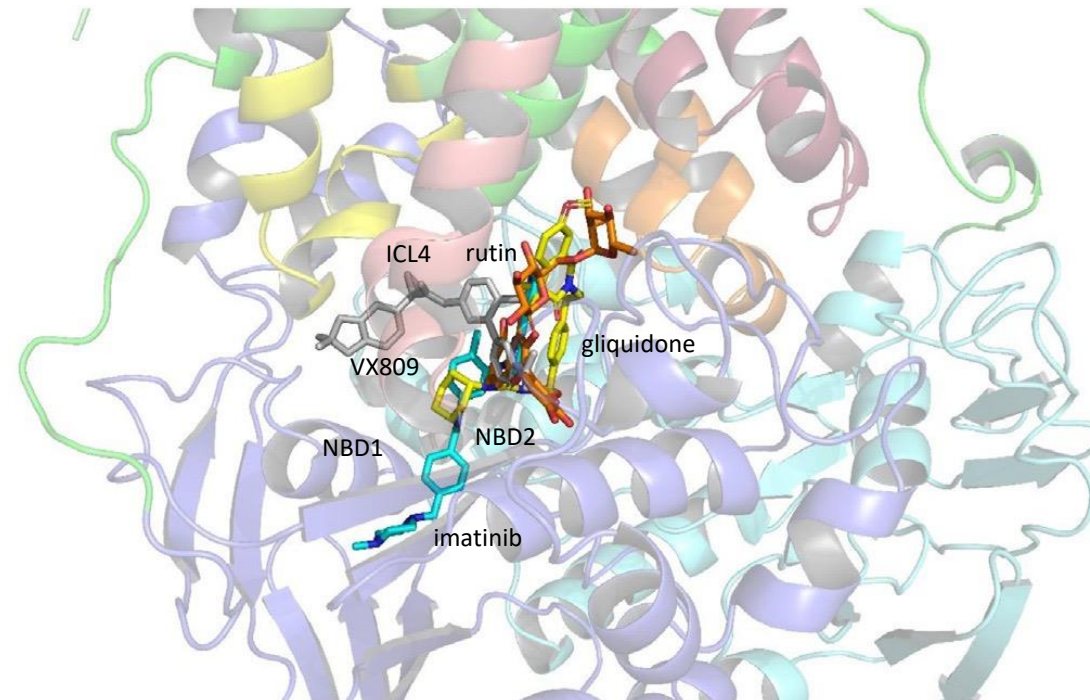
F508del-CFTR in cartoon: NBD1 in slate blue, NBD2 in cyan, ICL1 in yellow, ICL2 in orange, ICL3 in raspberry red, ICL4 in salmon pink, and TMs in green.

Three typologies of drugs can be defined depending on the location of residues involved in the binding:
- **VX809** , doxorubicin, nilotinib, aprepitant, bexarotene, eltrombopag, Zafirlukast and **rutin** bind the residues of the **NBD1** and **NBD2** and **ICL4**,
- **imatinib**, tadalafil and telmisartan bind the residues of **NBD1** and **ICL4**,
- **gliquidone** interacts only with the apical region of **NBD1**.
The binding energy ranking of the repositioned drugs was combined with their clinical assessment, and only **telmisartan**, **tadalafil**, and **rutin** were chosen for further computational and experimental evaluation.

Complex stability

RMSF evaluates how much a residue moves during a simulation. In addition, higher RMSF values indicate that the protein has more flexible domains.
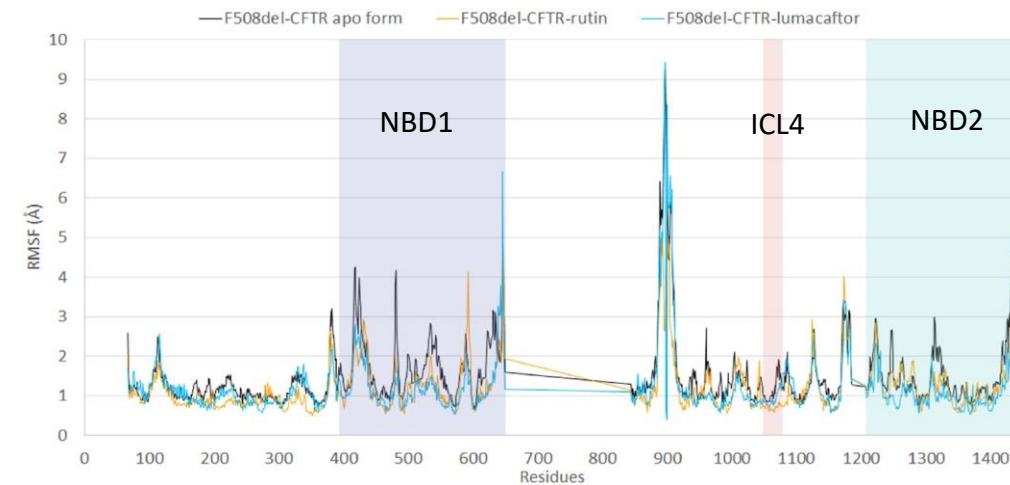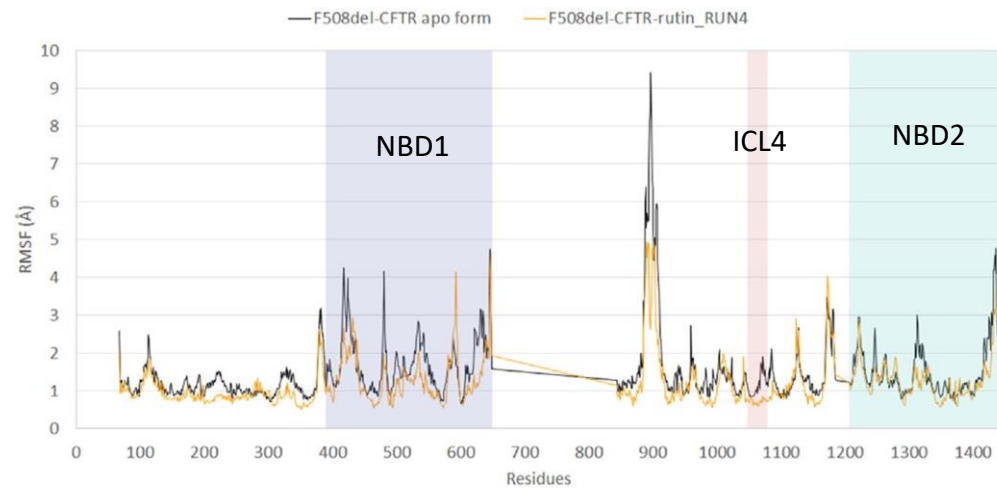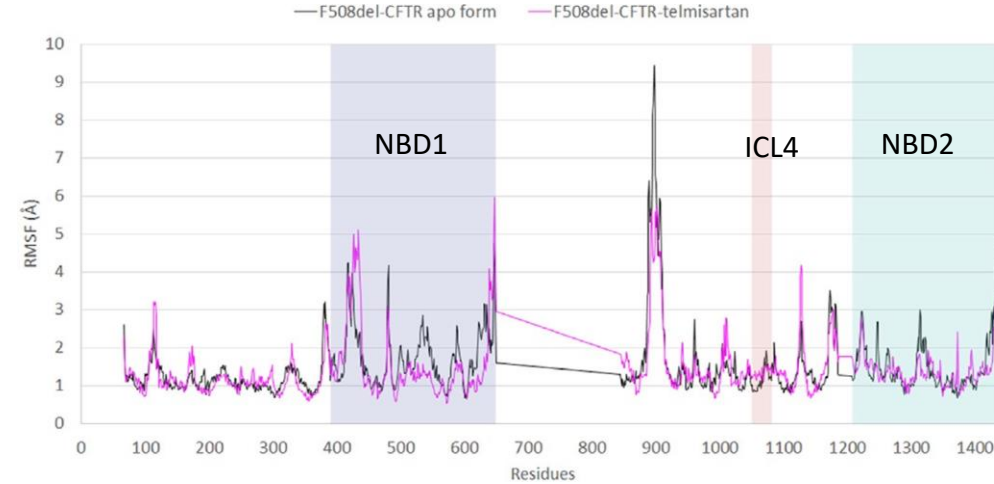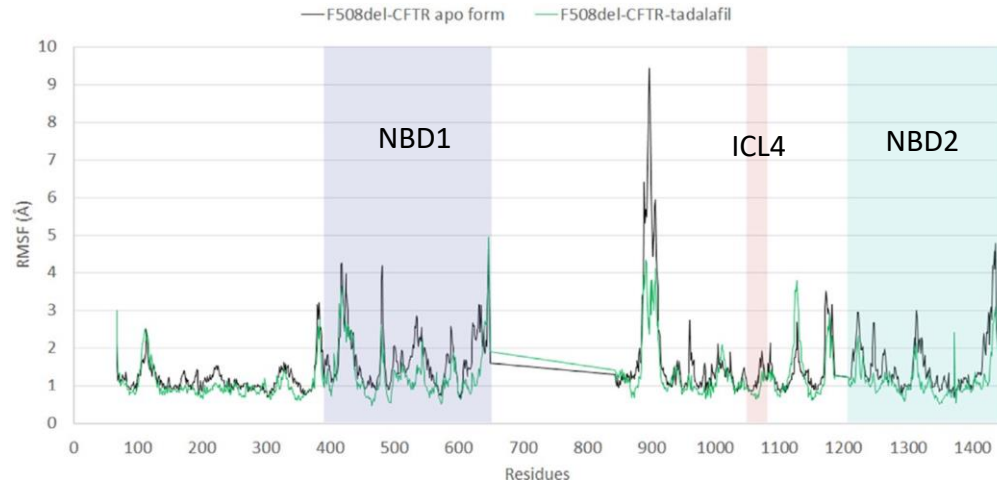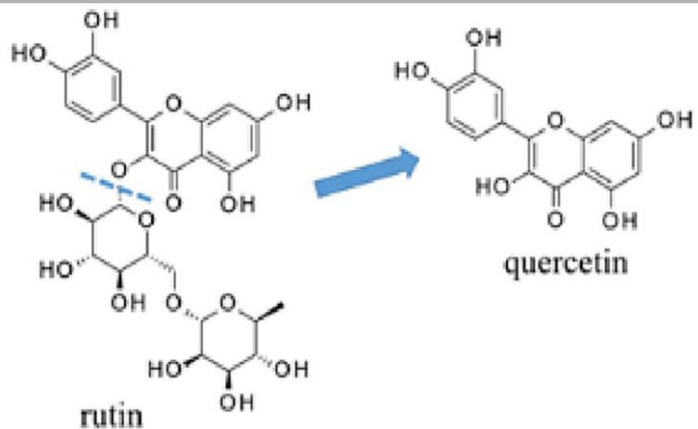

Tadalafil, telmisartan, and rutin show a lower or equal fluctuation than apo F508del-CFTR. Tadalafil, telmisartan, and rutin stabilize part of the NBD1 domain, while part of NBD2 is stabilised only by rutin and tadalafil. Rutin is the only compound able to stabilize the ICL4 domain, while telmisartan produces a higher ICL4 fluctuation when compared to the apo form.
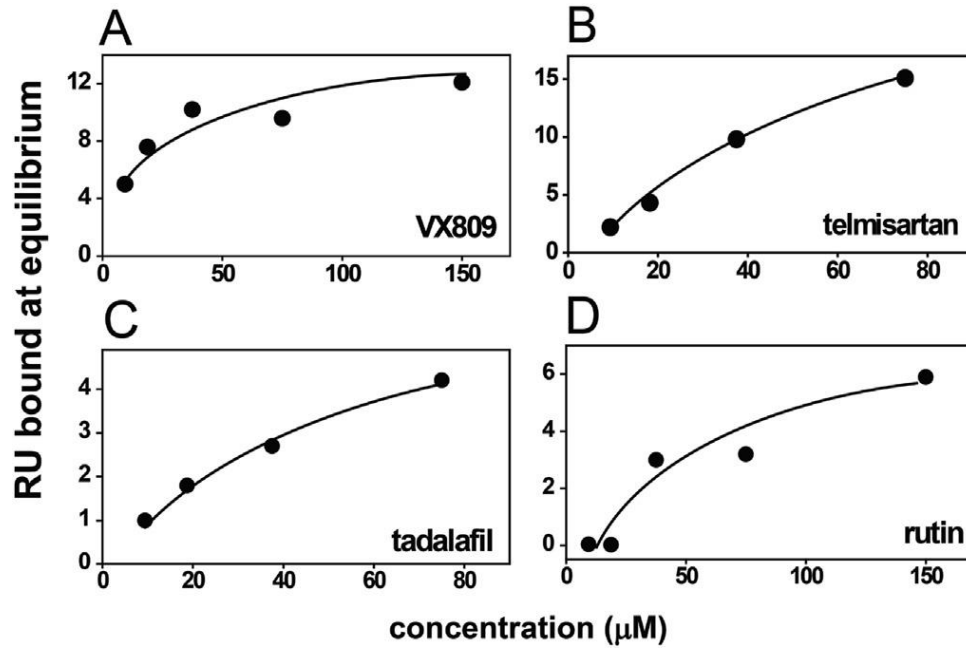
SPR analysis of the binding of the
repositioned drugs to F508del-CFTR

Binding of the compounds to sensorchip-immobilized F508del-CFTR. Kd values are reported in μM ± S.E.M. The number of repeated independent calculations are indicated in brackets.

Compound	Kd (μM)
VX809 (template)	47.8 18.36 (5)
telmisartan	189.3 ± 48.5 (3)
tadalafil	175.0 ± 50.4 (4)
Rutin	65.8 ± 27.3 (7)
quercetin	25.6 ± 10.2 (3)

Computational and biochemical validation of quercetin/ F508del-CFTR interaction

Rutin is bio-converted in the quercitin by human gut bacteria
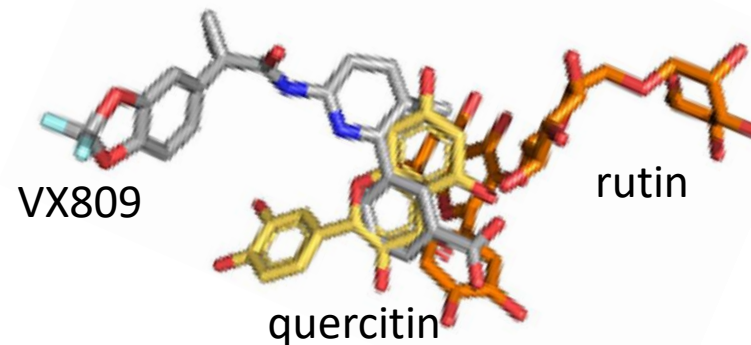
Overlapping of the binding poses of VX809, rutin and quercitin
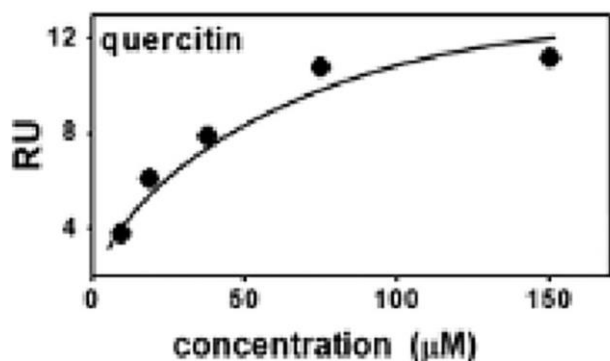
Binding of the compounds to sensorchip-immobilized F508del-CFTR. Kd values are reported in µM ± S.E.M. The number of repeated independent calculations are indicated in brackets.

| Compound	Kd (µM)
VX809 (template)	47.8 18.36 (5)
telmisartan	189.3 ± 48.5 (3)
tadalafil	175.0 ± 50.4 (4)
Rutin	65.8 ± 27.3 (7)
quercetin	25.6 ± 10.2 (3)
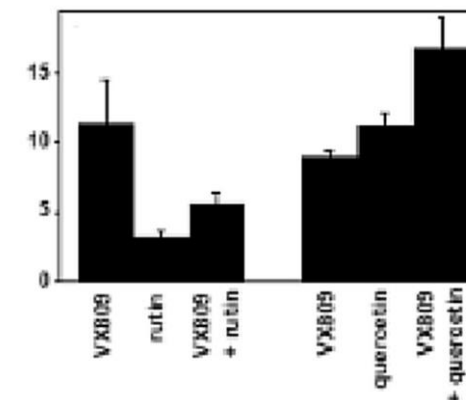
Quercitin binds F508del-CFTR in a dose-dependent and saturable manner with an affinity that is comparable to VX809, as demonstrated by SPR analysis. Competition binding assay shows that rutin partially inhibits VX809 to F508del-CFTR, while quercitin shows an additive effect due to its small dimension
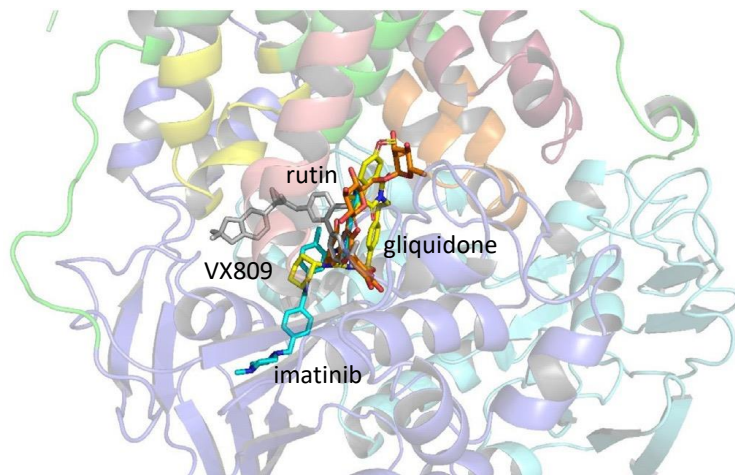
Three sub-regions around the aromatic acid part VX809 are not occupied by the lumafactor. Other drugs could occupy these sub-regions, which could enhance the effect of the corrector VX809.

Fpocket analysis evaluated the druggability score of the sub-regions.
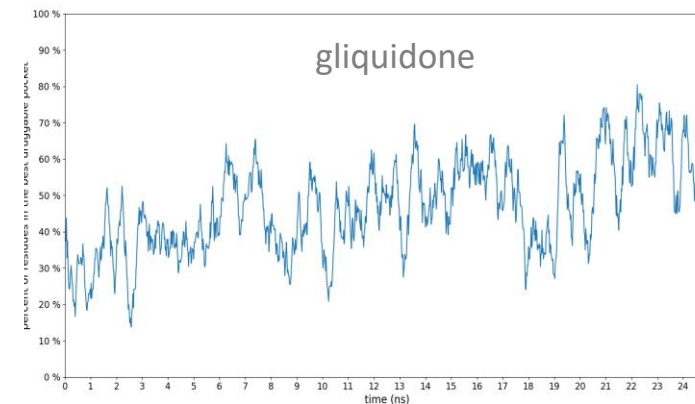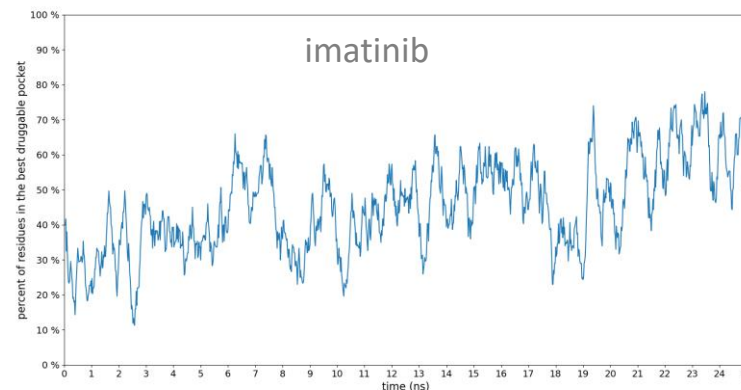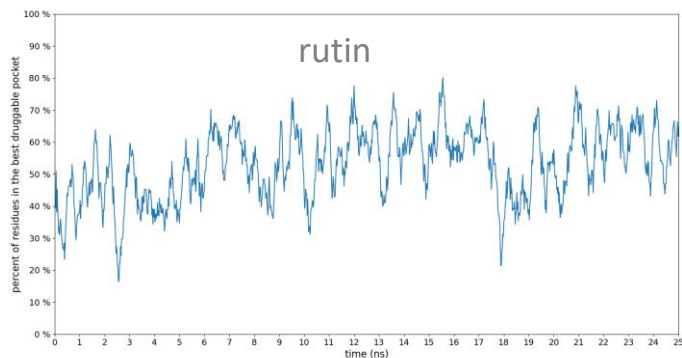

The plots show that the residues surrounding imatinib, rutin and gliquidone are endowed with high druggability scores along the molecular dynamics simulations. For these residues, a druggability score of more than 70% is observed.
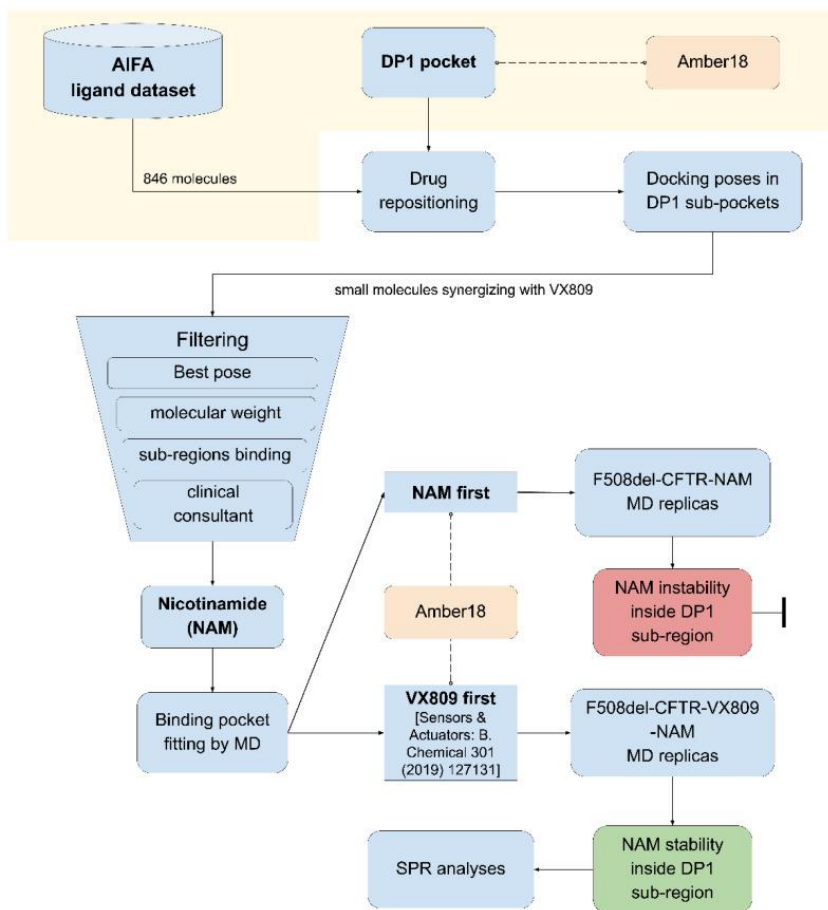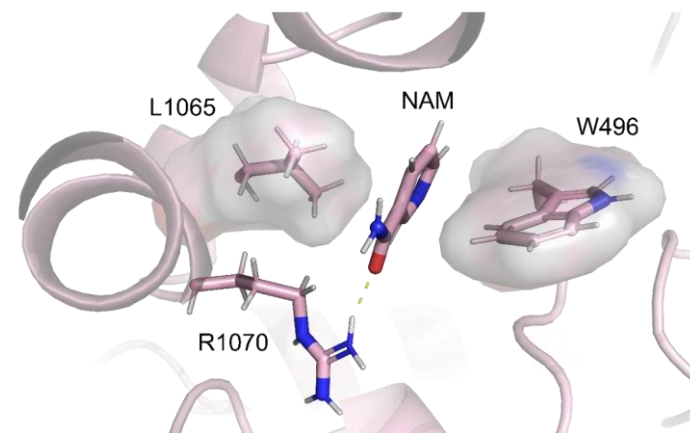
The docking results were filtered based on their binding into the druggable DP1 sub-regions and ranked based on binding energy value and MW.

Complex stability
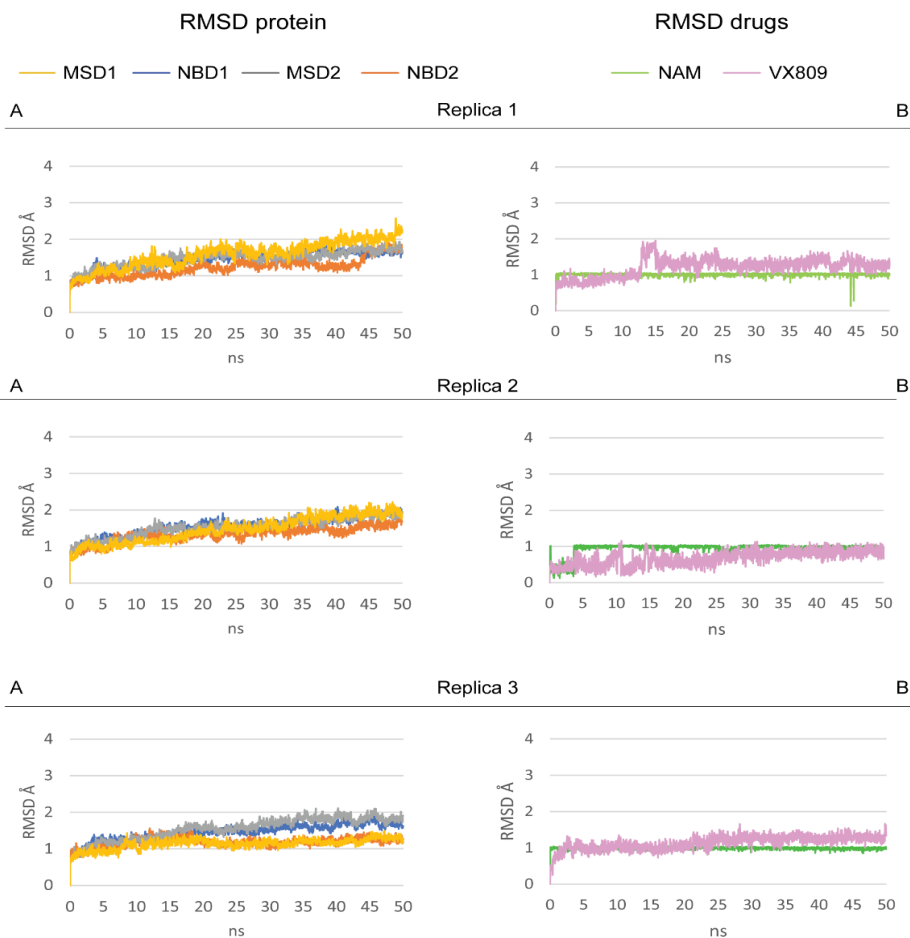
RMSD analysis of F508del-CFTR four domains (A), and NAM and lumacaftor (B) for the three replicas.


NAM contributes to generate contacts with W496 of NBD1 and L1065 and R1070 of ICL4

NAM binding pose in the F508del-CFTR-lumacaftor-NAM complex. The H-bond is shown in yellow dotted lines. Residues involved in hydrophobic interactions with NAM are depicted with their hydrophobic surfaces.

Experimental validation
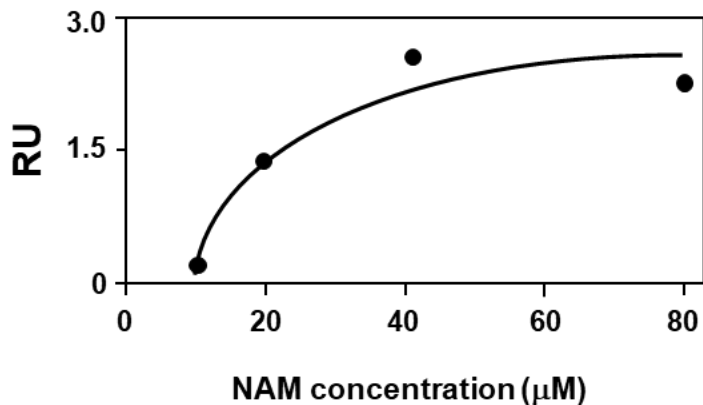
Surface plasmon resonance binding analysis confirmed that NAM effectively binds to F508del-CFTR
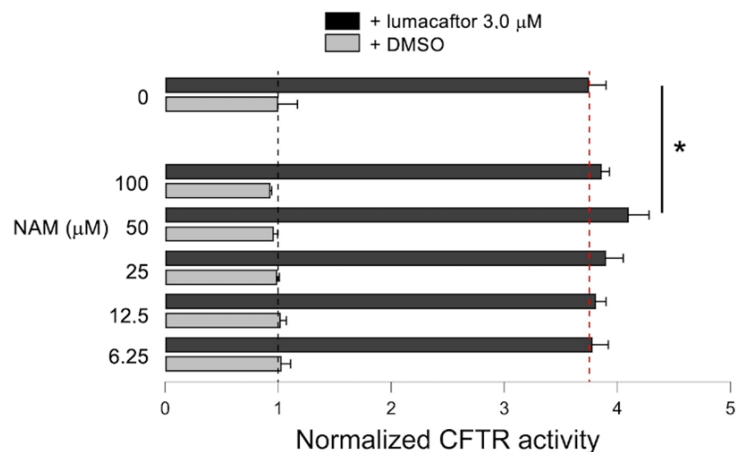
*Effect of NAM-lumacaftor co-treatment on mutant F508del-CFTR rescue. The bar graphs show F508del-CFTR activity in CFBE41o-cells stably expressing the HS-YFP. CFTR activity was determined as a function of the YFP quenching rate following iodide influx in cells treated for 24 h with DMSO in the absence (vehicle) or in the presence of lumacaftor (3.0 M) as a single agent or combined with the indicated concentrations of NAM. * p < 0.05.*