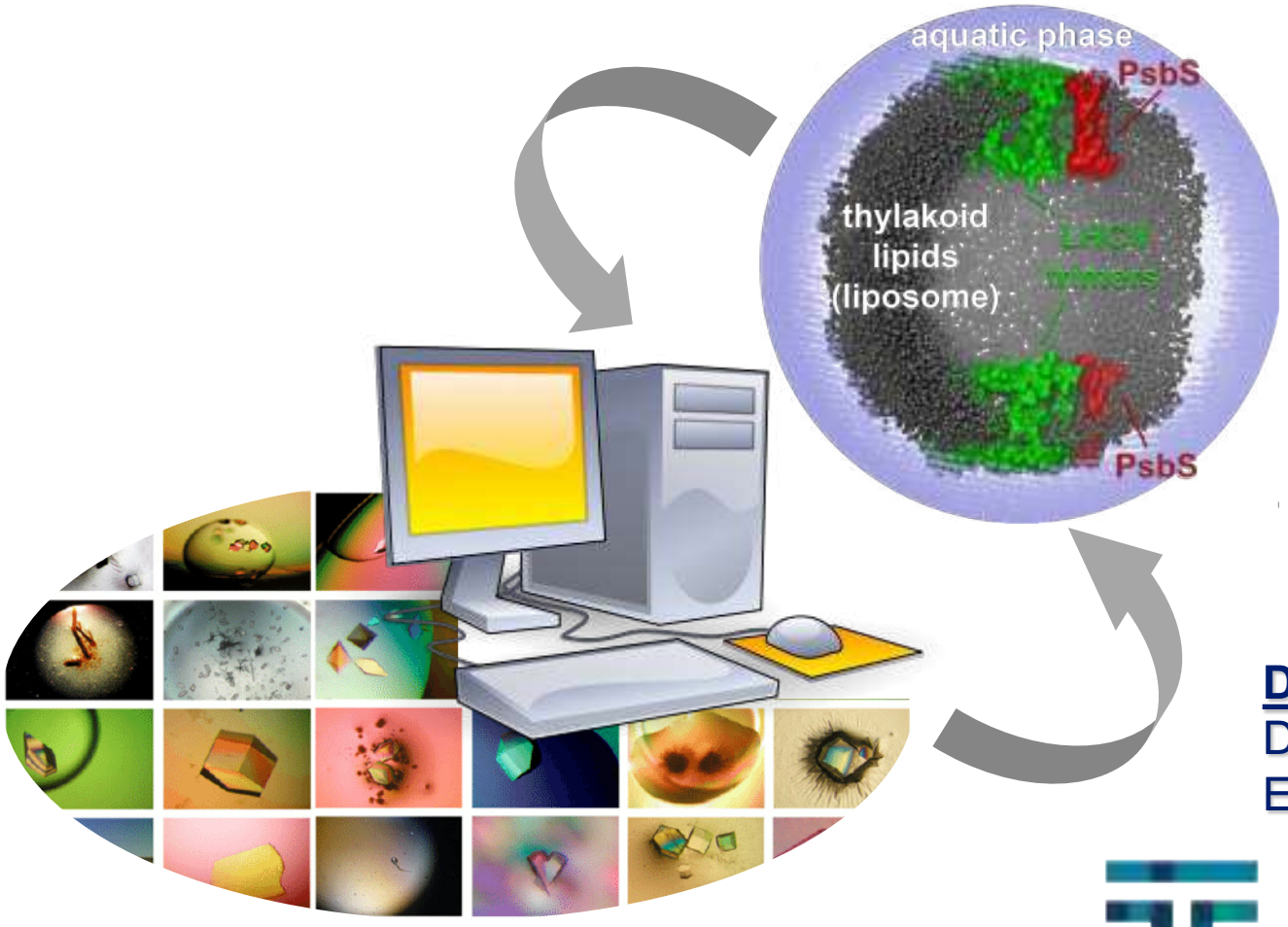
# High Performance Computing in Structural Biology

***...from rigid crystals to molecular dynamics***

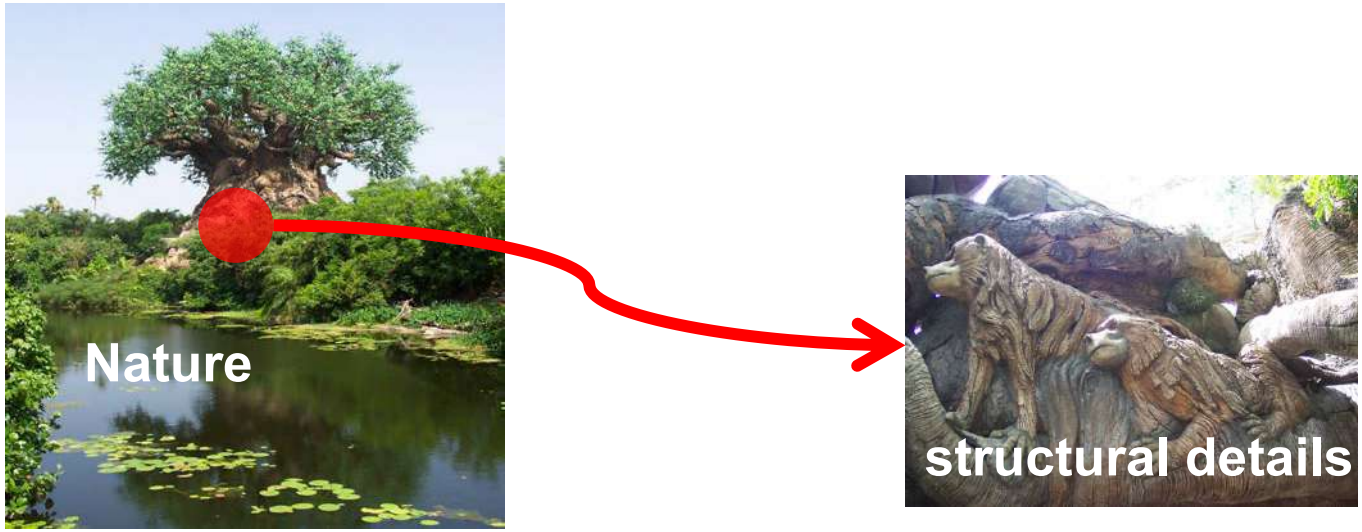



**Dr. Vangelis Daskalakis**
Department of Chemical Engineering

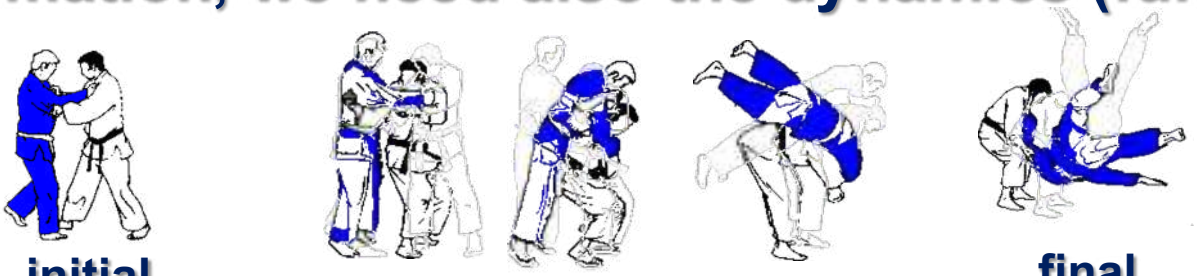
Cyprus University of Technology

# Protein Structure and Dynamics

➔Structure is the basic element to understand nature




➔Not all information is stored within a structure. To find a final conformation, we need also the dynamics (function)



initial conformation

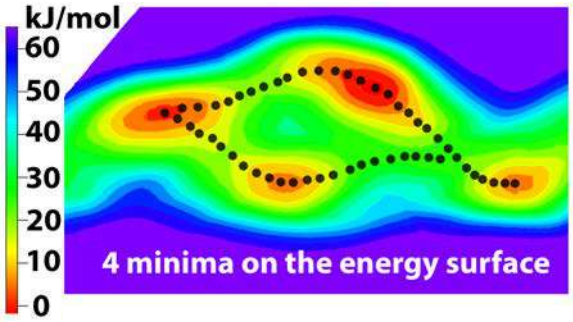
dynamics

final conformation

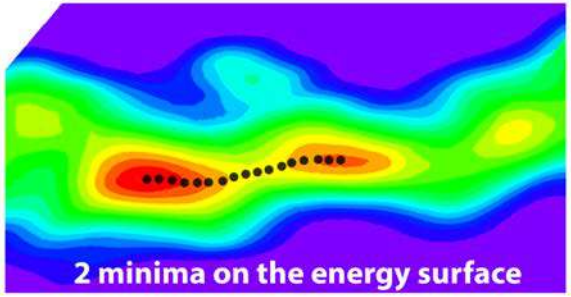
# Methodology - Approach

**Protein Structure and Function: the motion of atoms and molecules – protein atoms 'move' on free energy surfaces between energy minima and convert energy for biological work (bioenergetics)**

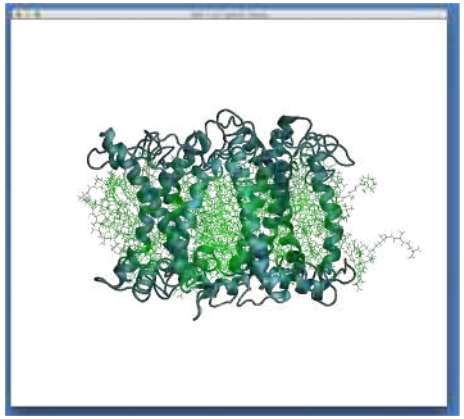
**Flexible Protein**

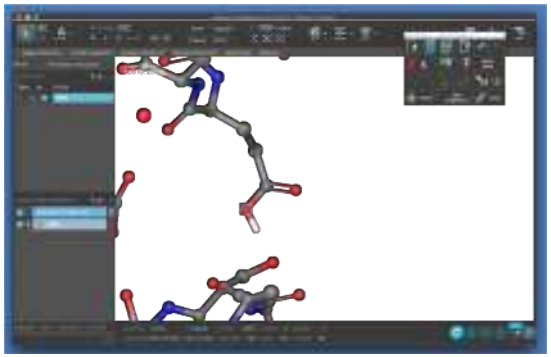



**Rigid Protein**

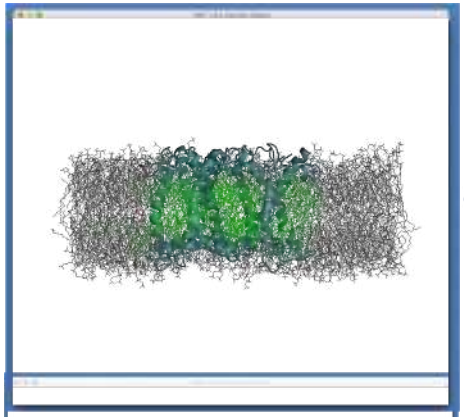



# Structure: The basic element for research on bioenergetics

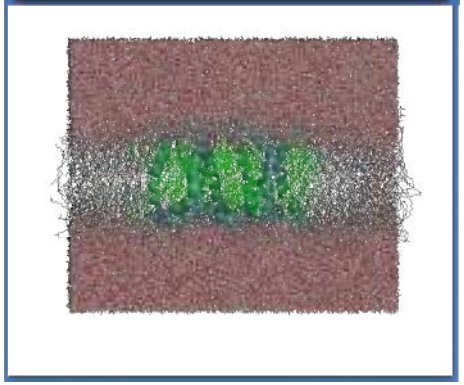


protonate



add membrane

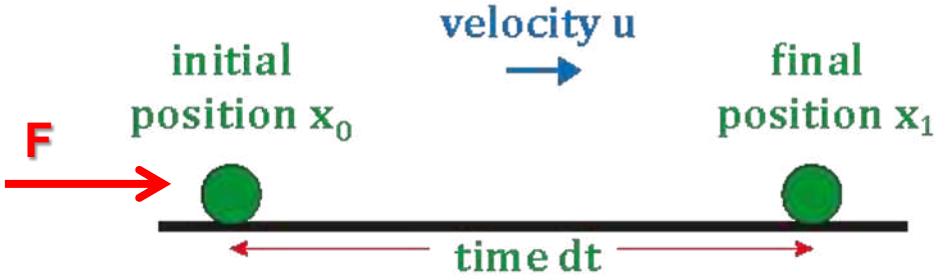
mutate



solvate

# Beyond the structure: Molecular Dynamics

**Molecular Dynamics: a simple theoretical basis**

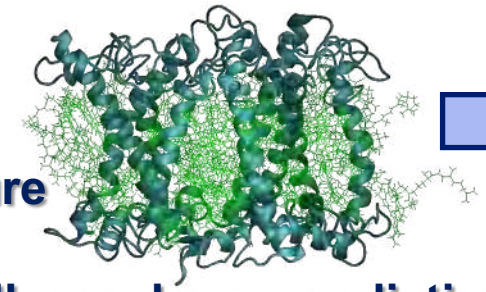



$$\mathrm{F} = m * a \rightarrow F = m \times \frac{du}{dt} \rightarrow F = m \times \frac{d^2x}{dt^2}$$

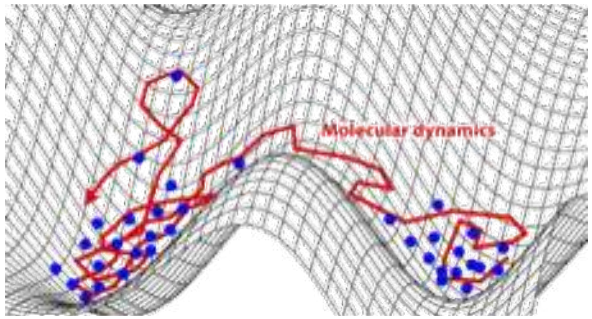
**That is how we also move atoms in proteins!**

# Methodology: From Structure to Molecular Dynamics

Initial Structure



(crystallography or prediction)




Free Energy Surface

Structure -2 Structure -1

Energy

force

$$F = \frac{-dEi}{dr_i} = mi\frac{d^2r_i}{dt^2}$$

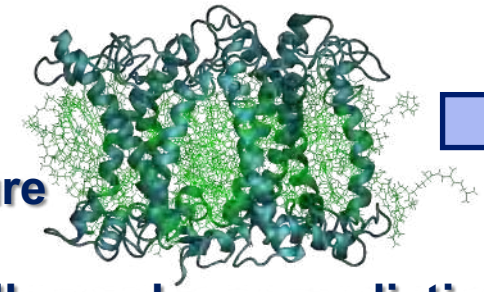
time

mass

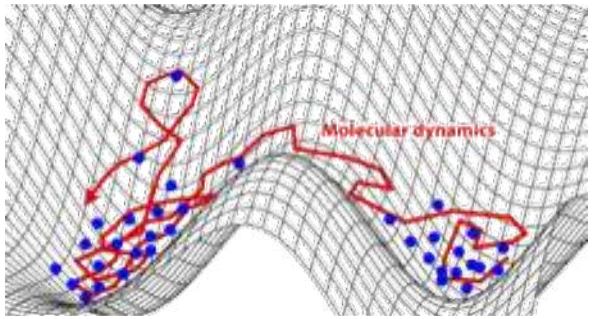
$\rightarrow r_i(t)\ trajectory\ of\ atoms$

# Methodology: From Structure to Molecular Dynamics

Initial Structure

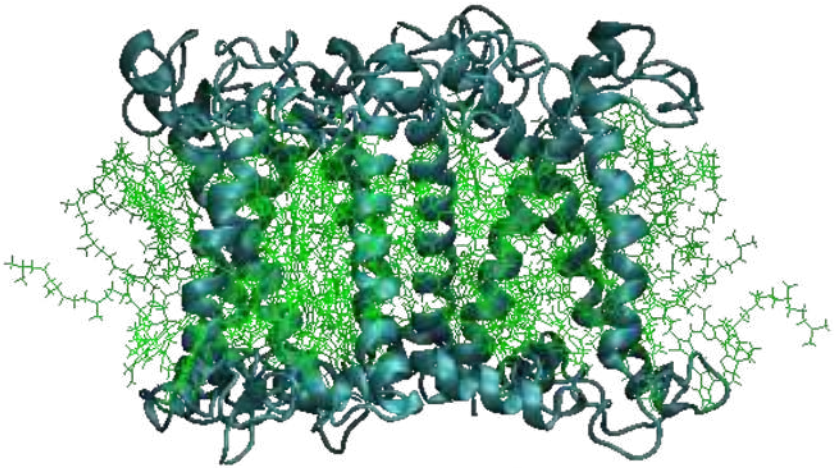


(crystallography or prediction)




Free Energy Surface

Structure -2 Structure -1



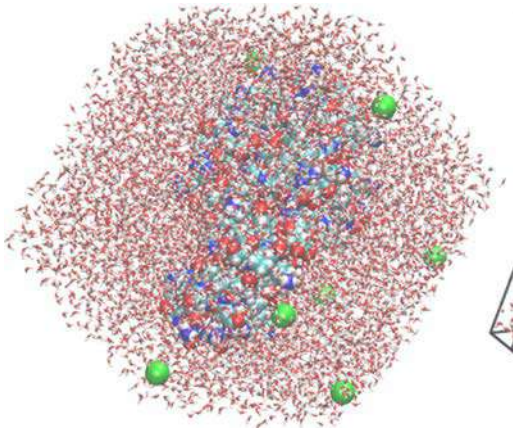
we observe protein 'function', like under a powerful microscope in correlation with experiments!

$$F = \frac{-dEi}{dr_i} = mi\frac{d^2r_i}{dt^2}$$

Force, Energy, mass, time

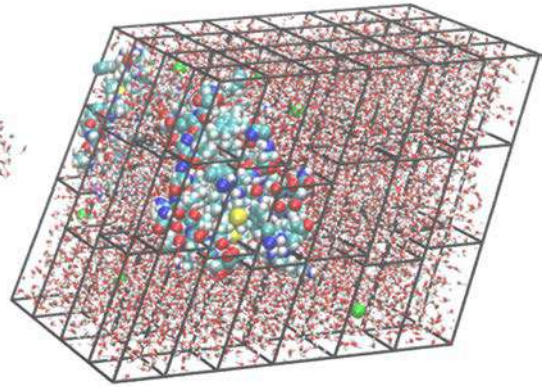
$\rightarrow r_i(t)\ trajectory\ of\ atoms$

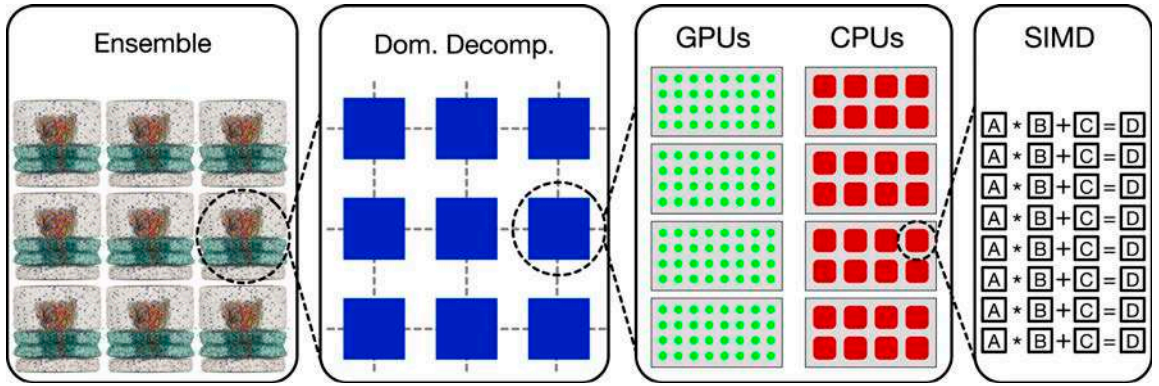
# Methodology – the HPC approach



simulation cell

## Domain Decomposition






**~185000 atoms/ simulation cell**
**256 std-cpu cores/ simulation cell**
**9 cell replicas ➔ 2304 std cpu cores/ run**

**SoftwareX 2015,** 1–2, 19-25

# Structural Biology meets Molecular Dynamics

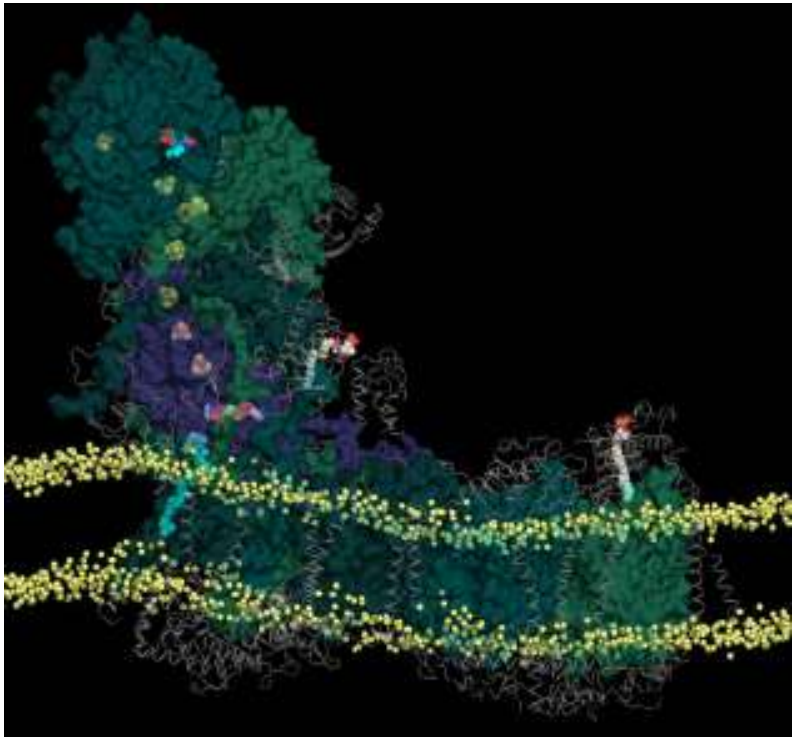
**High resolution structure and molecular simulations of a key bioenergetic protein**

RESEARCH ARTICLE | STRUCTURAL BIOLOGY

## High-resolution cryo-EM structures of respiratory complex I: Mechanism, assembly, and disease

Kristian Parey[1,2,3], Outi Haapanen[4], Vivek Sharma[4,5], Harald Köfeler[6], Thomas Züllig[6], ...
+ See all authors and affiliations

*Science Advances* 11 Dec 2019:
Vol. 5, no. 12, eaax9484
DOI: 10.1126/sciadv.aax9484

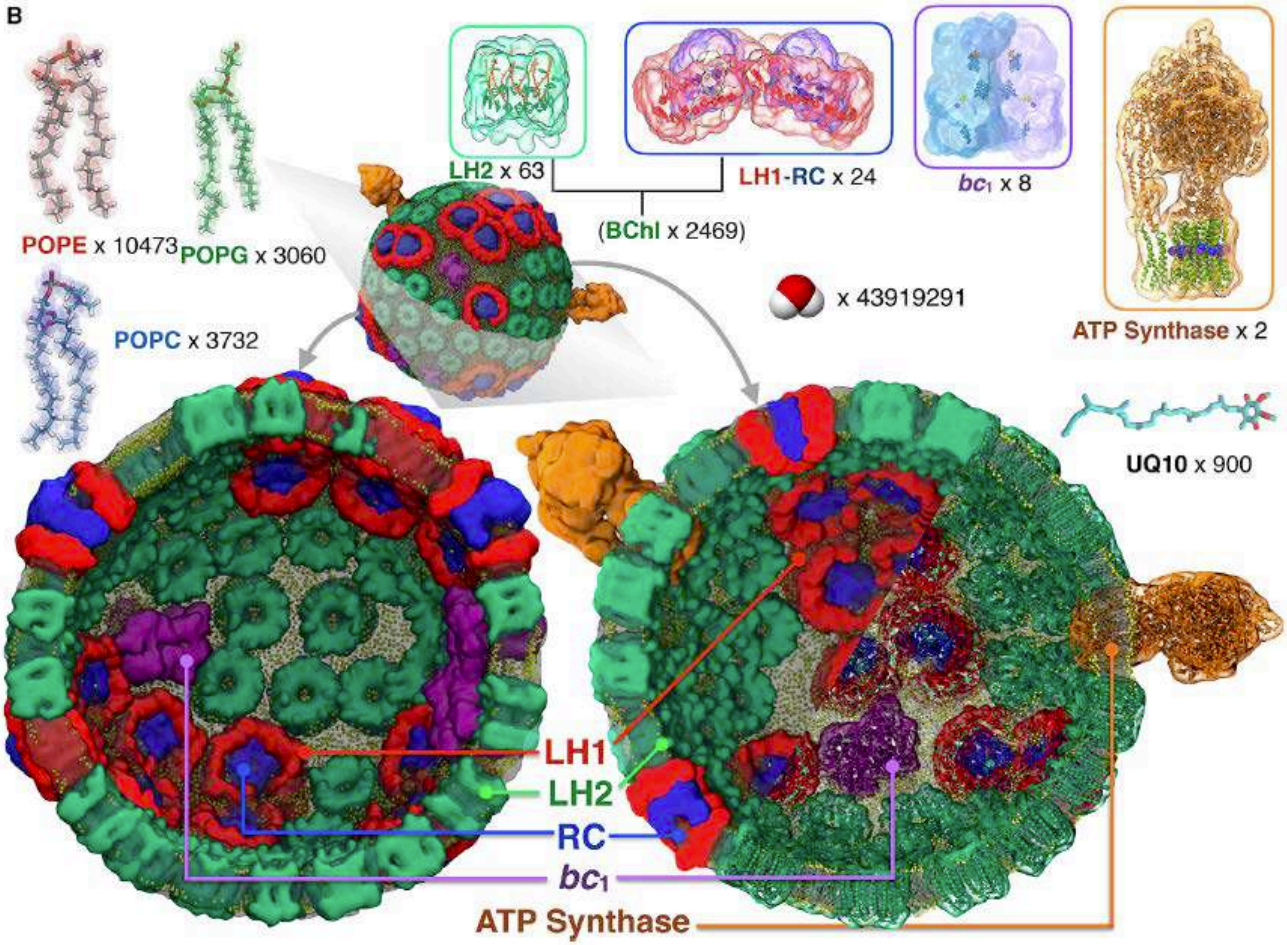


# Structural Biology meets Molecular Dynamics

## Atoms to Phenotypes: Molecular Design Principles of Cellular Energy Metabolism

Cell

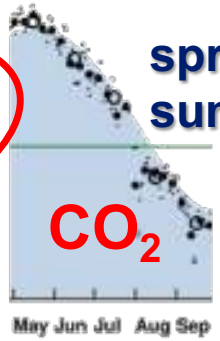
Volume 179, Issue 5, 14 November 2019, Pages 1098-1111.e23

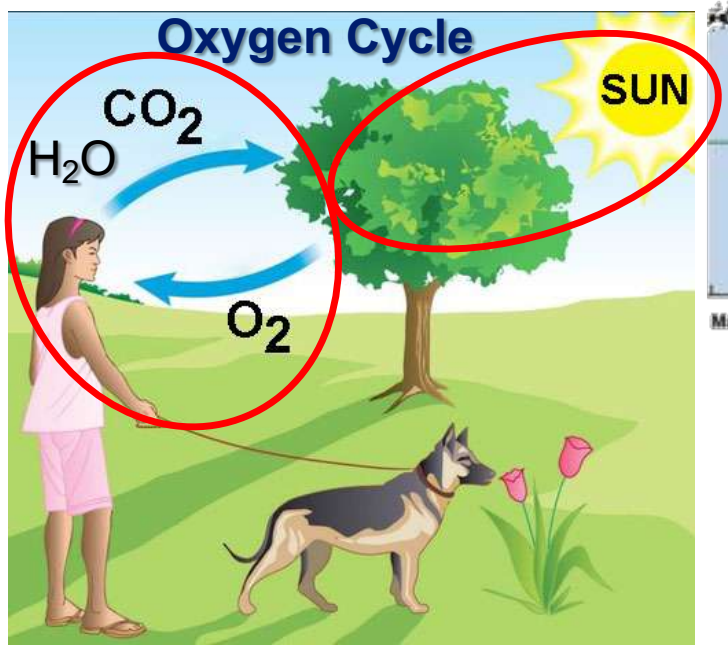



**Simulation of an entire cell organelle: a Photosynthetic chromatophore vesicle from Purple bacteria at all atom resolution!**

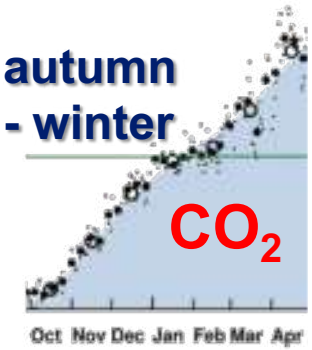
# Bioenergetics – Oxygen Cycle

**Photosynthesis**
**Oxygen evolution**




**Cellular Respiration**
**Oxygen Reduction**



# Bioenergetics 2020

***An active field of research*: cellular respiration or photosynthesis, are essential to most aspects of cellular metabolism, therefore to life itself: nature.com/ sciencemag.org (2020)**

**Structural basis for energy transfer in a huge diatom PSI-FCPI supercomplex**

Caizhe Xu, Xiong Pi [...] Jian-Ren Shen

*Nature Communications* **11** , 1–12

**Essential role of accessory subunit LYRM6 in the mechanism of mitochondrial complex I**

Etienne Galemou Yoga, Kristian Parey [...] Heike Angerer

*Nature Communications* **11** , 1–8

**Structural basis for energy transfer in a huge diatom PSI-FCPI supercomplex**

Caizhe Xu, Xiong Pi [...] Jian-Ren Shen

*Nature Communications* **11** , 1–12

**Thermodynamic efficiency, reversibility, and degree of coupling in energy conservation by the mitochondrial respiratory chain**

Mårten Wikström & Roger Springett

*Communications Biology* **3** , 1–9

**Charge transfer from the carotenoid can quench chlorophyll excitation in antenna complexes of plants**

Lorenzo Cupellini, Dario Calvani [...] Benedetta Mennucci

*Nature Communications* **11** , 1–8

**Ultrafast structural changes within a photosynthetic reaction centre**

Robert Dods, Petra Båth [...] Richard Neutze

*Nature* , 1–5

Light harvesting in oxygenic photosynthesis: Structural biology meets spectroscopy

Roberta Croce[1,*], Herbert van Amerongen[2]

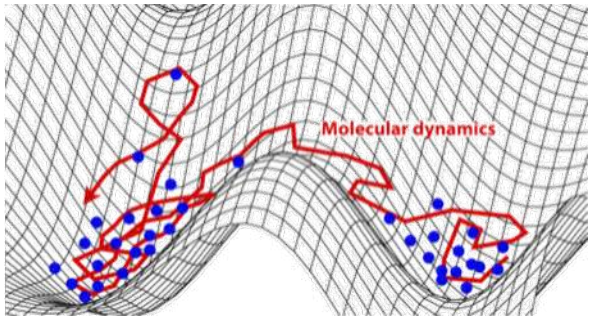
+ See all authors and affiliations

*Science* 21 Aug 2020:
Vol. 369, Issue 6506, eaay2058
DOI: 10.1126/science.aay2058

# Bioenergetics – Research Focus

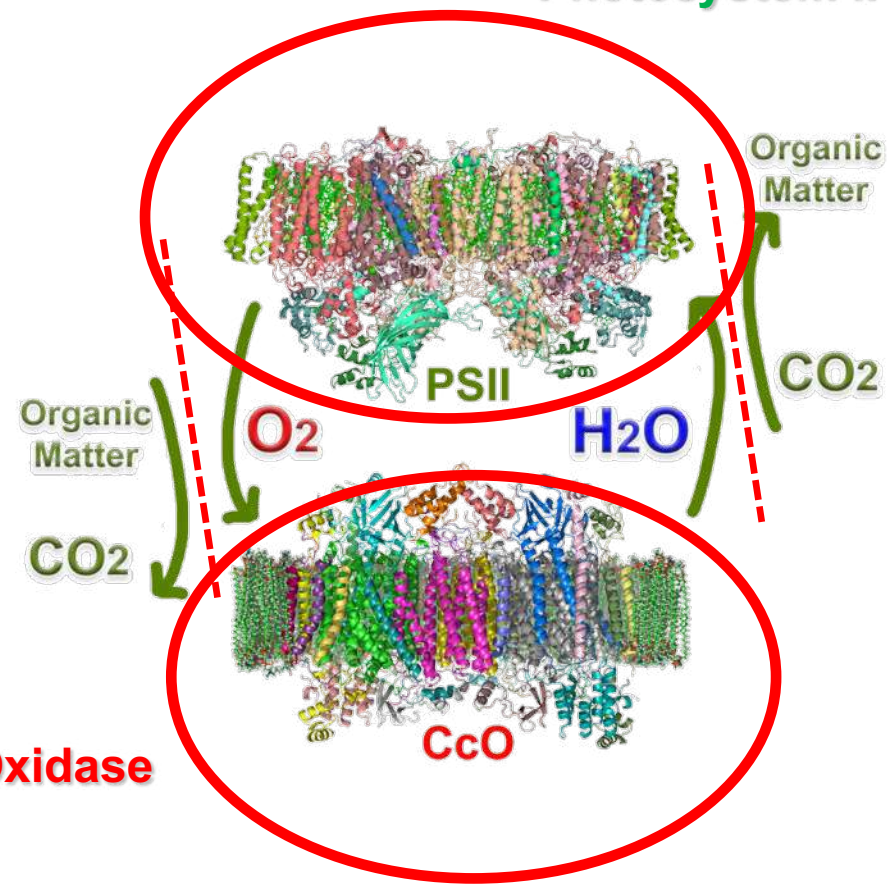
**photosynthesis and cellular respiration rely on protein complexes to harvest and convert energy**

**Structural details of the conversion are missing!**




**Structure -2** **Structure -1**

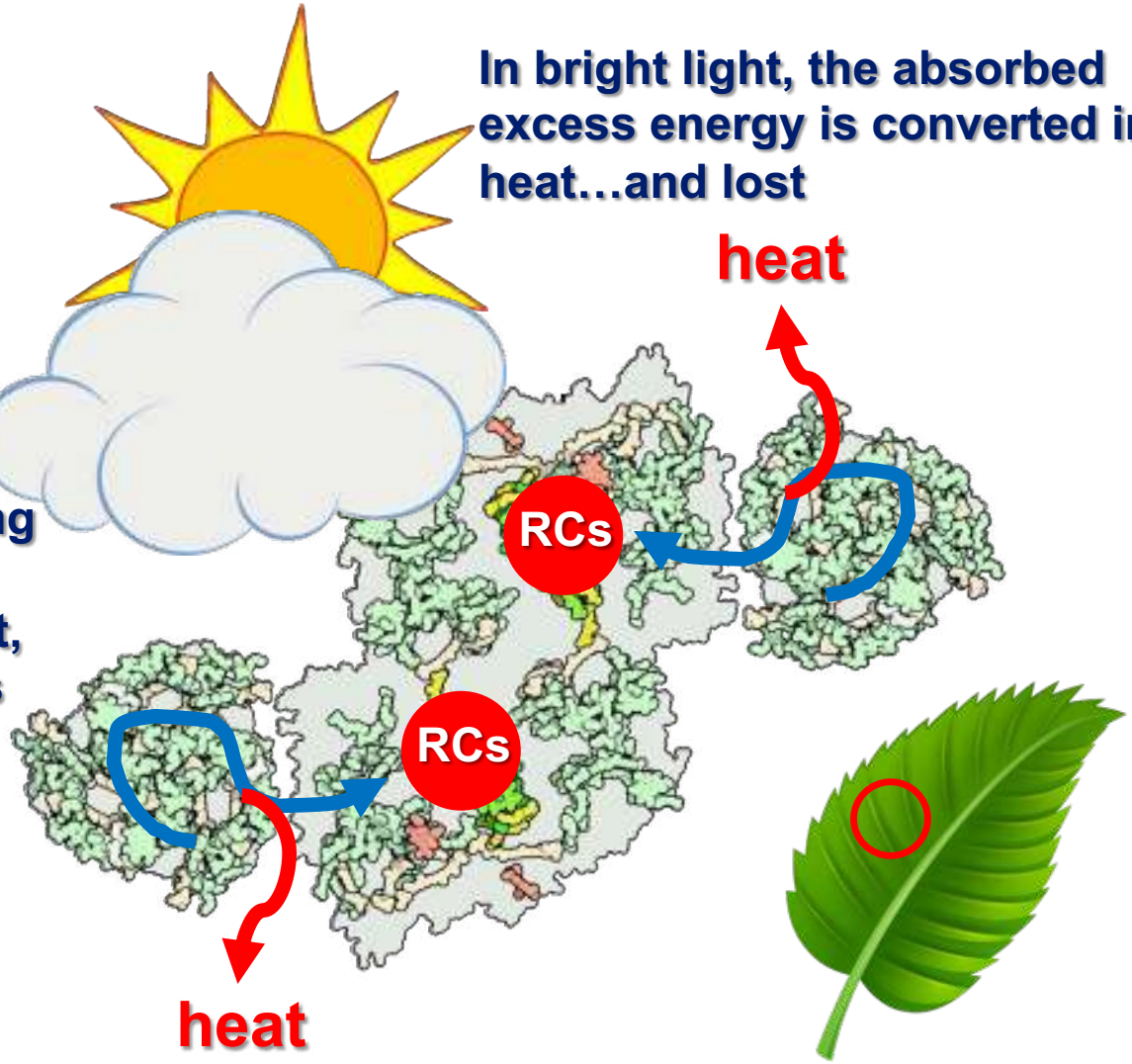
**Photosystem II**




**Cytochrome *c* Oxidase**

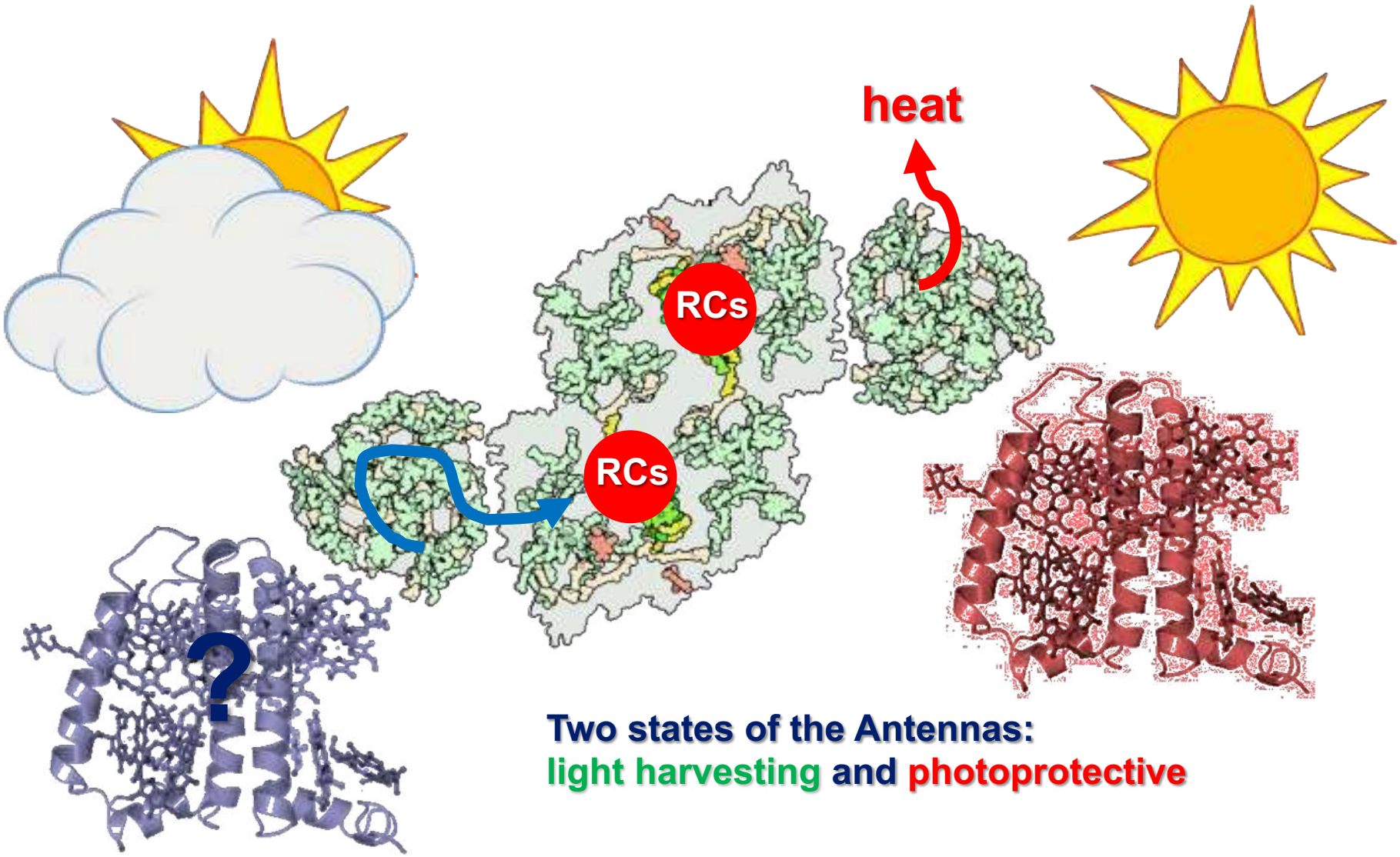
# Photosynthesis: a complicated story evolves

In bright light, the absorbed excess energy is converted into heat…and lost




Antennas or Light Harvesting Complexes (LHCII) around Photosystem II absorb Light, direct it to Reaction Centers

# Photosynthesis: a complicated story evolves

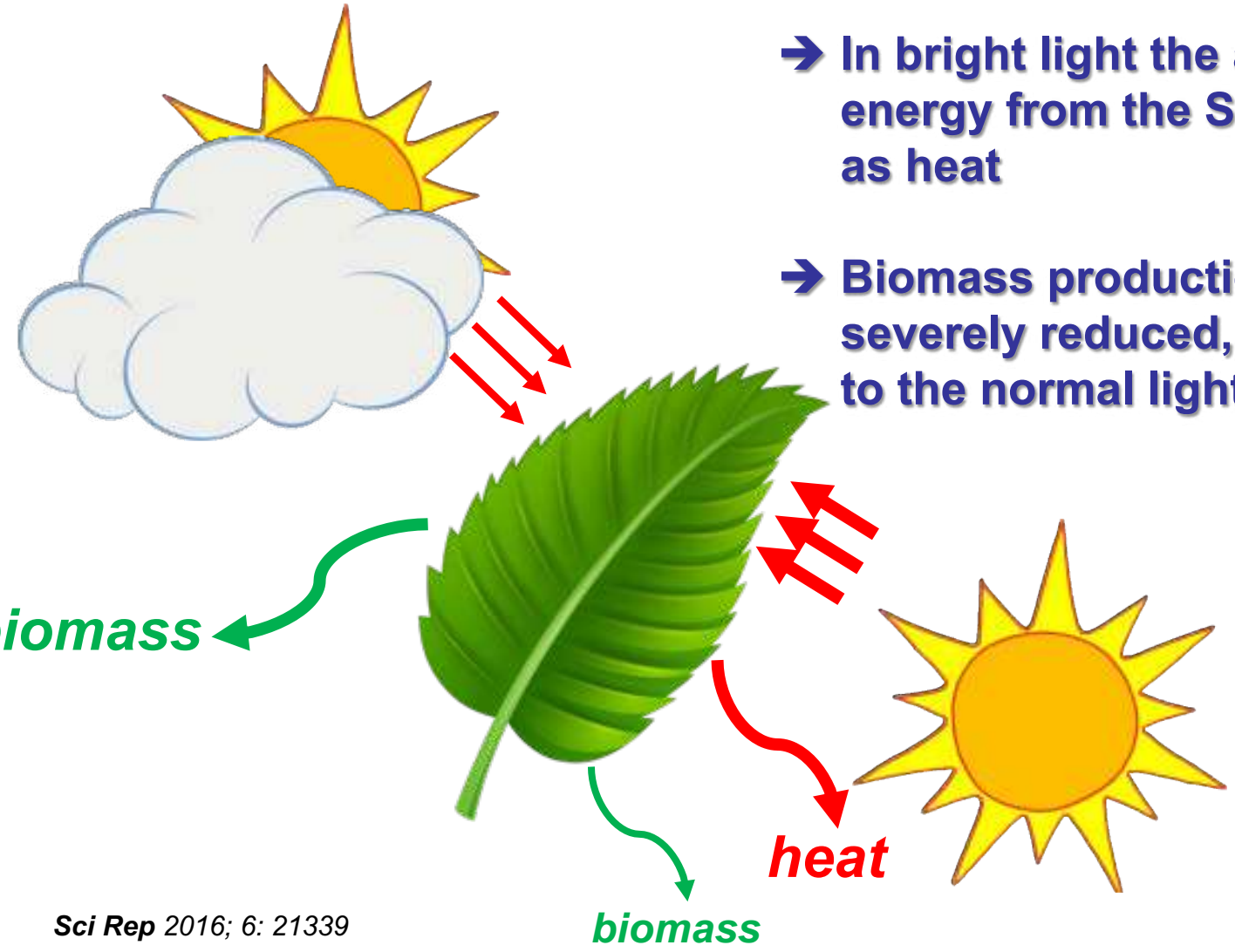



**Two states of the Antennas:**
**light harvesting and photoprotective**

**No crystal structure yet exists for the Light Harvesting conformation!**

# Photoprotection

*There is a global interest – Why ?*




➔ In bright light the absorbed energy from the Sun is lost as heat

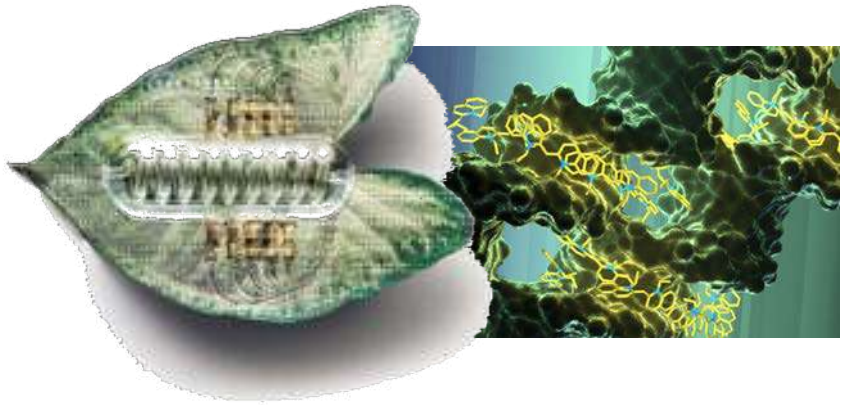
➔ Biomass production is severely reduced, compared to the normal light case

***Sci Rep*** *2016; 6: 21339*

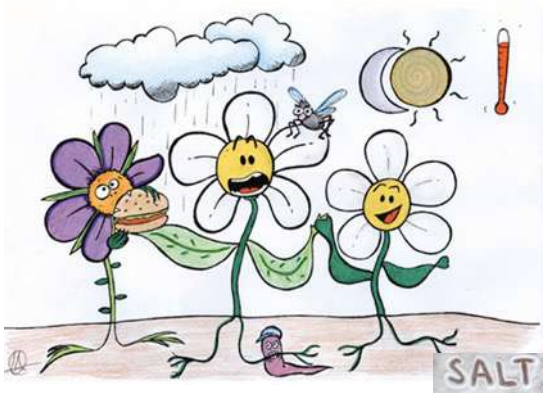
# Photoprotection

*There is a global interest – Why ?*

***Frontiers in plant science 2013**, 4:273*
***Science 2016**, 352(6290): 1210*
***Science 2016**, 354(6314): 857*
***Nature Communications 2020**, 11: 1*
***Science 2020**, 369 (6506): eaay2058*



➔**Development of artificial photosynthesis based on the light harvesting structure of LHCII**

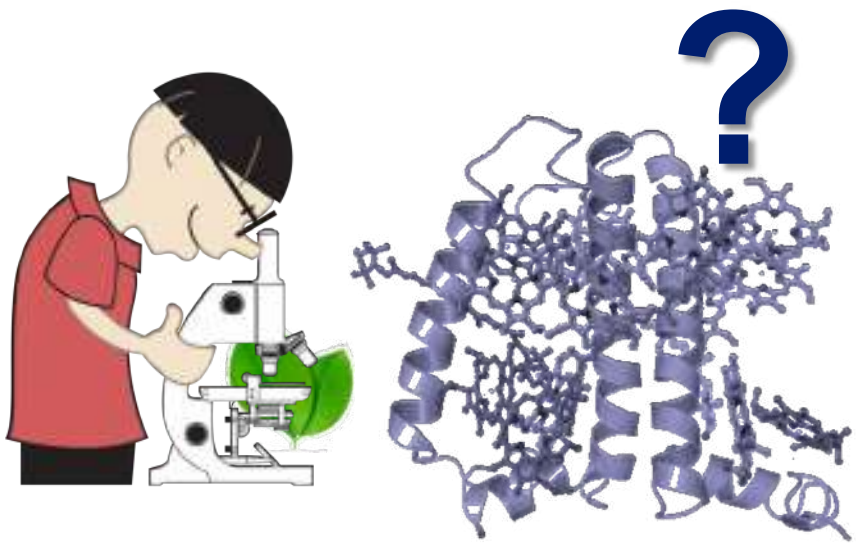


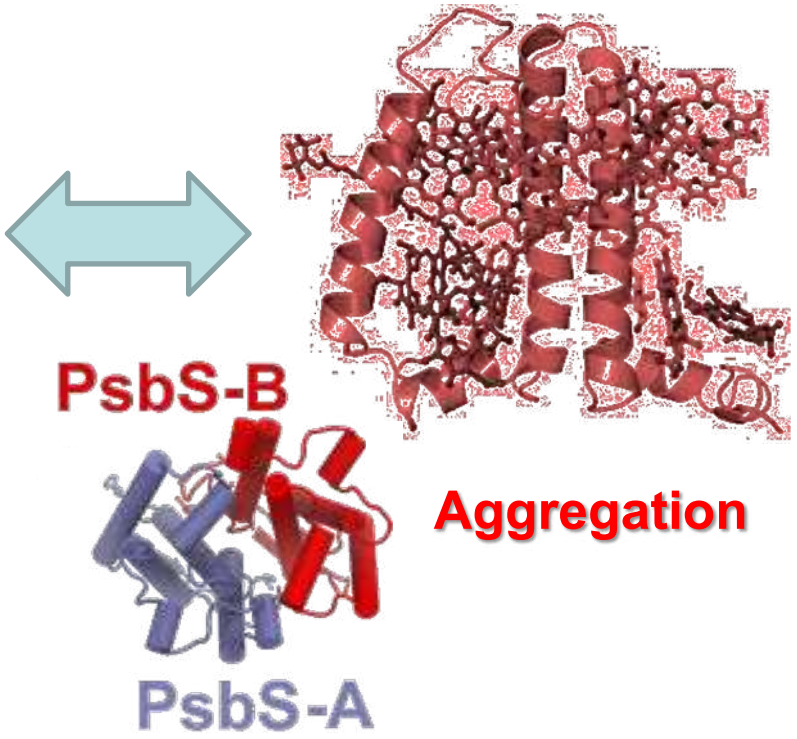

➔ **tolerant plants and increase of biomass (harsh conditions of soil, water, temperature)**

***Frontiers in plant science 2013**, 4:273*
***Science 2016**, 352(6290): 1210*
***Science 2016**, 354(6314): 857*

# Photoprotection

**First we need to resolve the Light Harvesting Conformation – Structure!**



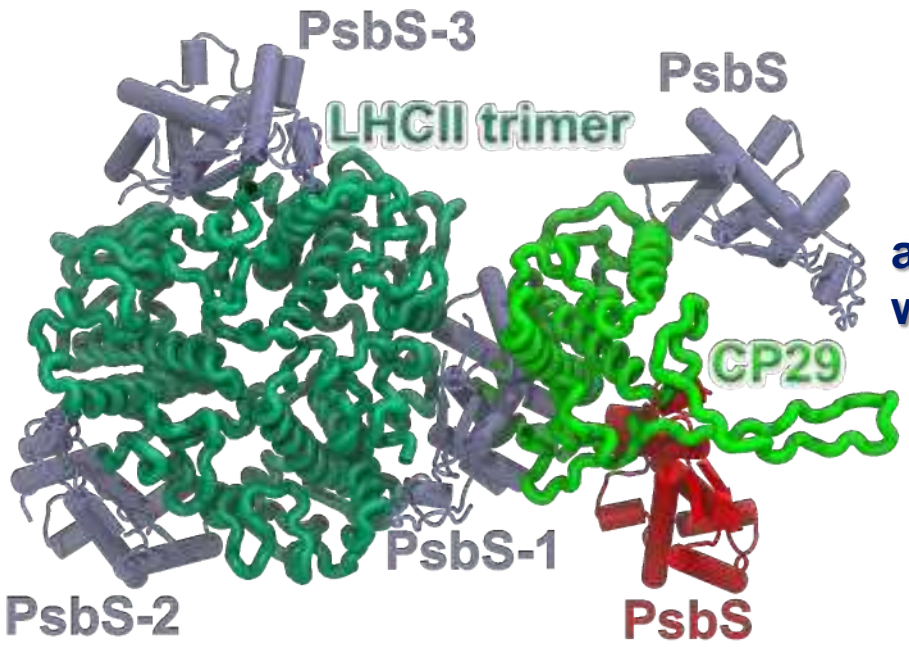




***In vivo the transition is achieved by:***

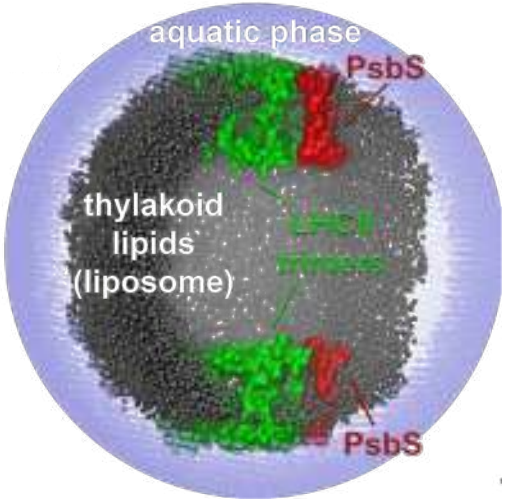
***what happens at the molecular-structural level?***

# Antenna Model

**Our model is the major LHCII trimer, embedded in a thylakoid membrane, or a liposome**



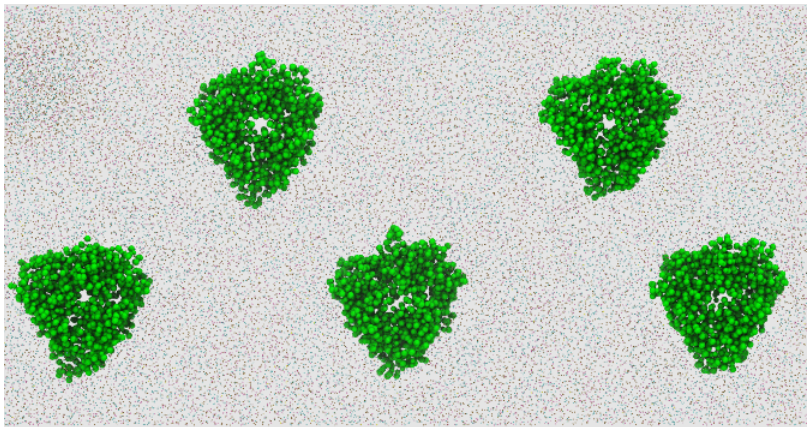

**accurate prediction of PsbS position within the available crystal structure**

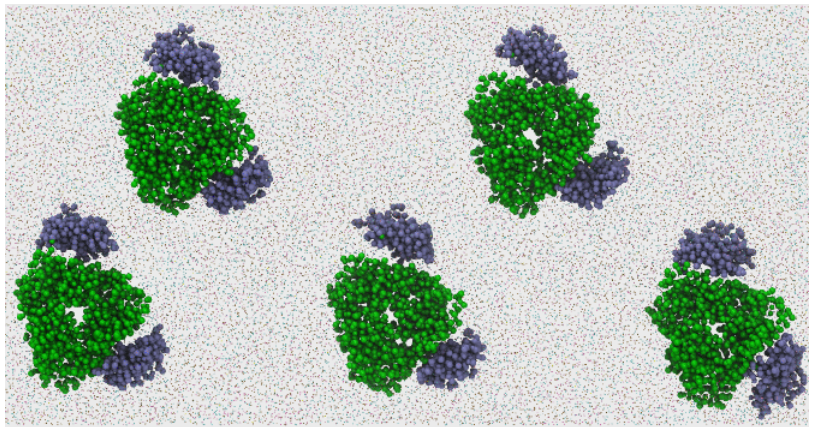



***Phys. Chem. Chem. Phys.* 2018**, 20, 11843 – 11855
***BBA-Biomembranes* 2019,** *1861*, 183059

# Protein Protein Associations

**LHCII aggregation in the absence and presence of PsbS, as we would watch it on a microscope**



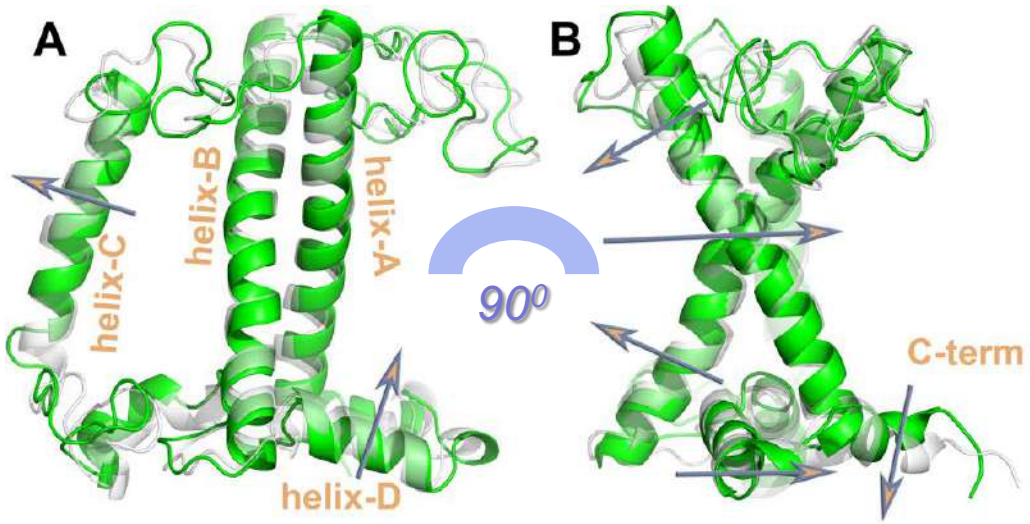


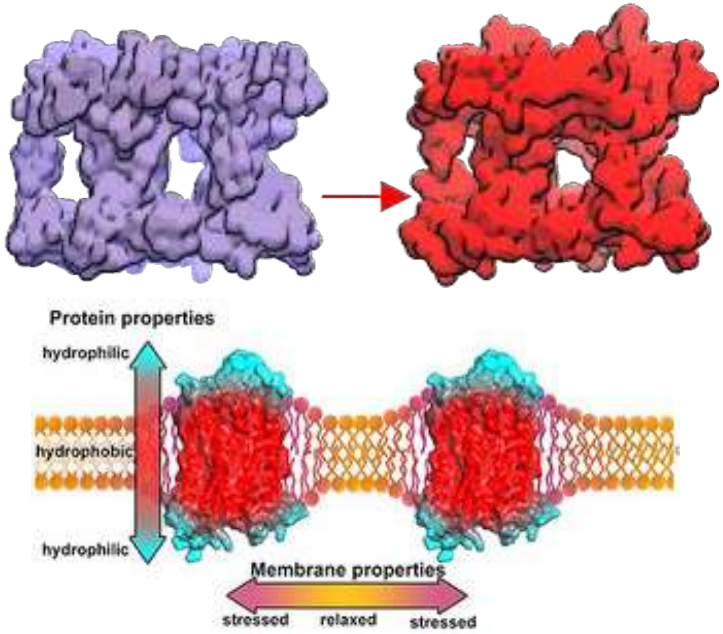
***Phys. Chem. Chem. Phys. 2018***, 20, 11843 – 11855
***BBA-Biomembranes* 2019,** *1861*, 183059

# Protein Protein Associations

**We have monitored the shape changes in LHCII, that also induce changes in the thickness of the thylakoid membrane under photoprotection (experimental confirmation)**

*sampling duration:150μs, literature state-of-the-art ~14 μs*

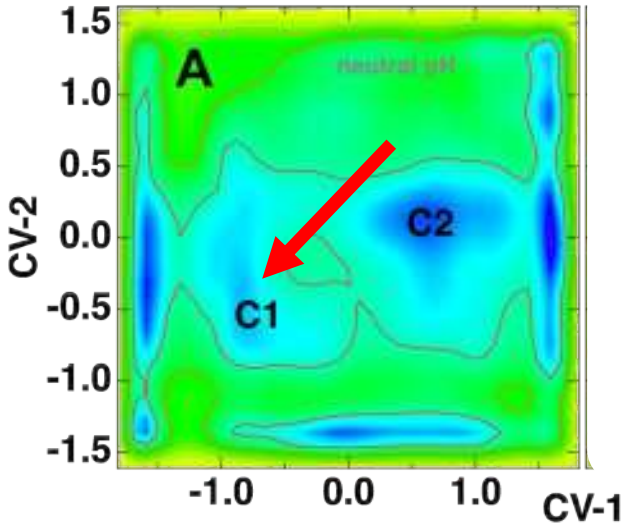


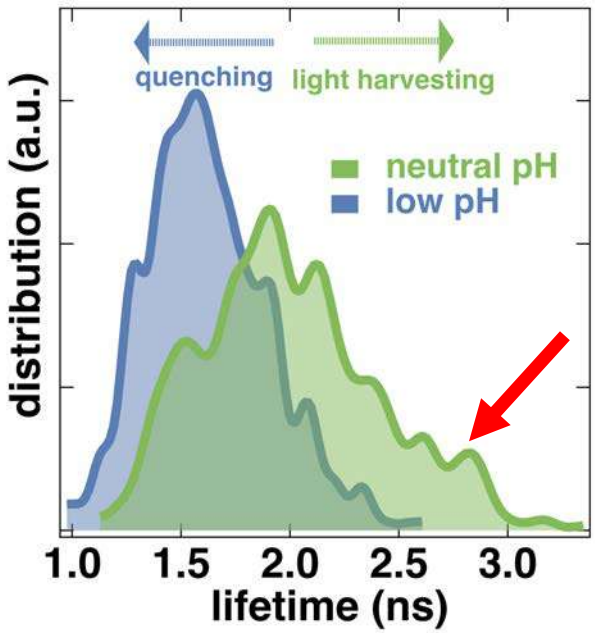



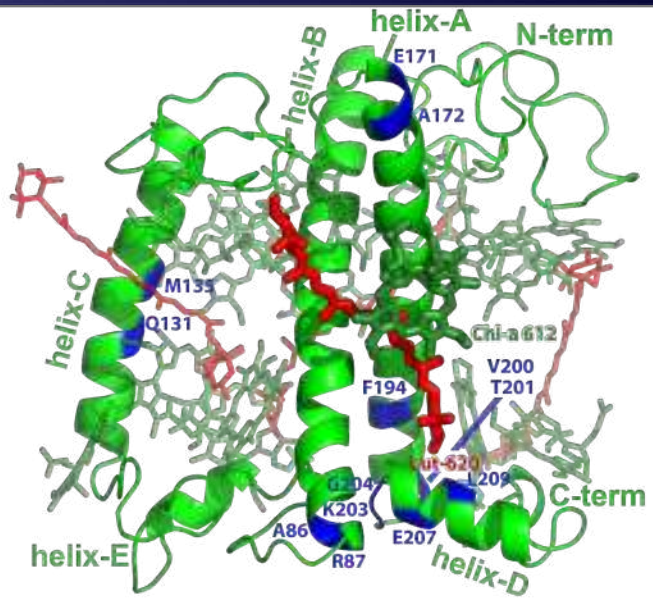

**Light Harvesting Conformation → Photoprotection**


***J. Chem. Phys.* 2019**, *141, 214105*
***J. Phys. Chem. B*, 2019** 123, 45, 9609-9615
***Chem .Commun. 2020 (RSC)**, 56:11215*

# The Light Harvesting State




**we have identified the elusive Light Harvesting LHCII structure with large Chlorophyl excited state lifetime (not dissipating energy as heat!)**

**➔ targeted experimental studies (mutations)**


***J. Phys. Chem. B*, 2019** 123, 45, 9609-9615
***Chem .Commun. 2020 (RSC)**, 56:11215*

# Thank you!


**PROJECT: 18–2018194641 “CDynLHCII””– 17.000.000 std cpu core hrs (12 months)**

**SuperMUC-NG**
consists of 6,336 Thin compute nodes each with 48 cores and 96 GB memory
144 Fat compute nodes each 48 cores and 768 GB memory per node

In total **311,040 compute cores** with a main memory of 719 TB and a peak performance of 26.9 PetaFlop/s are available.

# SARS-CoV-2 The Team

Cyprus University of Technology


**Cyprus University of Technology**
**Dept. Chemical Engineering**
**Assist. Prof. Vangelis Daskalakis, Ph.D**
Computational Physics – Biophysics Lab


UNIVERSITY OF CRETE


**University of Crete**
**School of Medicine**
**Prof. Elias Castanas MD, PhD**
**Mr. Athanassios A. Panagiotopoulos, Ph.D Candidate**
Dept. of Experimental Endocrinology

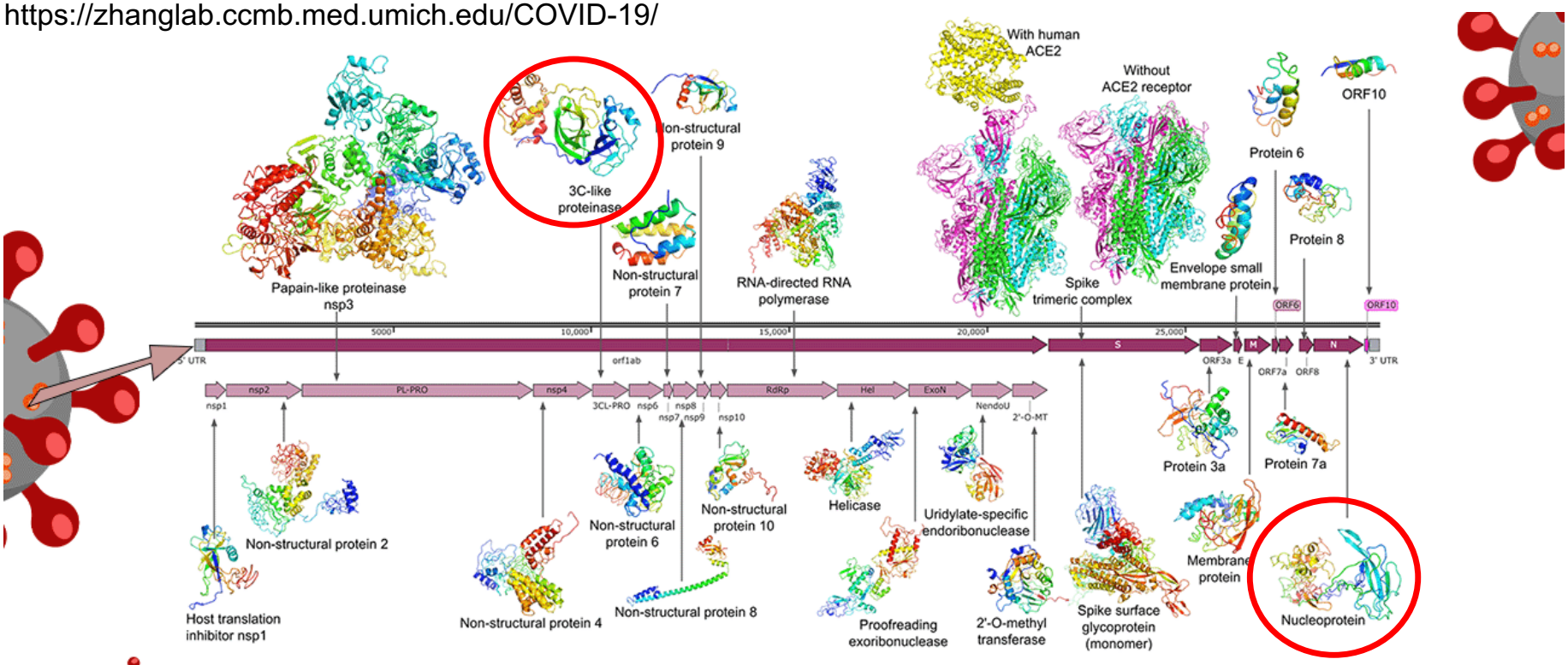
**Prof. George Sourvinos, PhD**
Dept. of Clinical Virology

**Mrs. Danai – Maria Kotzampasi**, MSc Candidate

**University of Crete**
**Dept. Biology**
**Prof. Stergios Pirintsos, Ph.D**
Biology of Organisms, Populations, Environmental and Marine Biology, Natural Products.

# The targets

https://zhanglab.ccmb.med.umich.edu/COVID-19/




**Genome-wide structure and function modeling of SARS-COV-2**

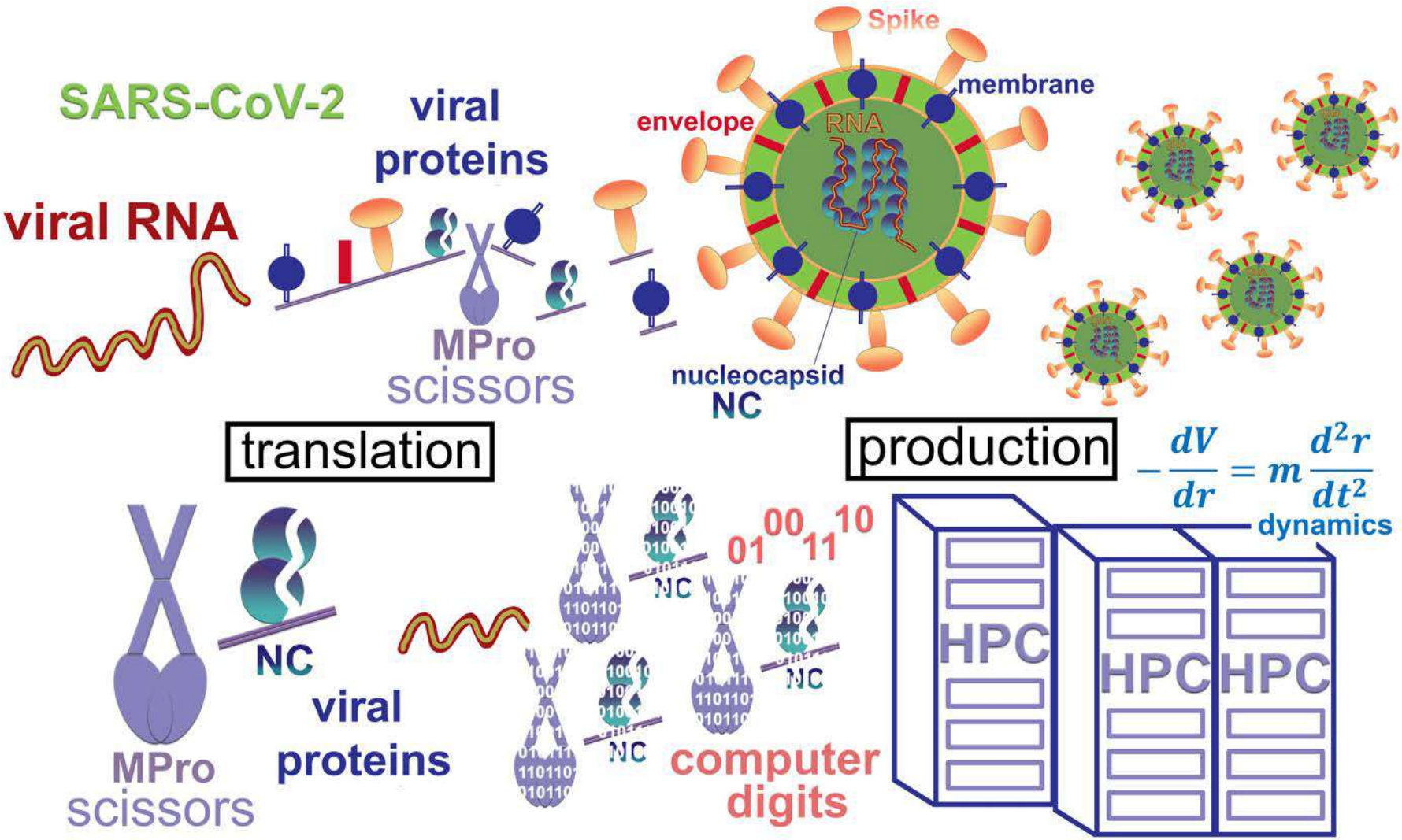
**3CL-Pro**

**cleaves an inactive polyprotein into 11 functional viral proteins… (active as a dimer)**

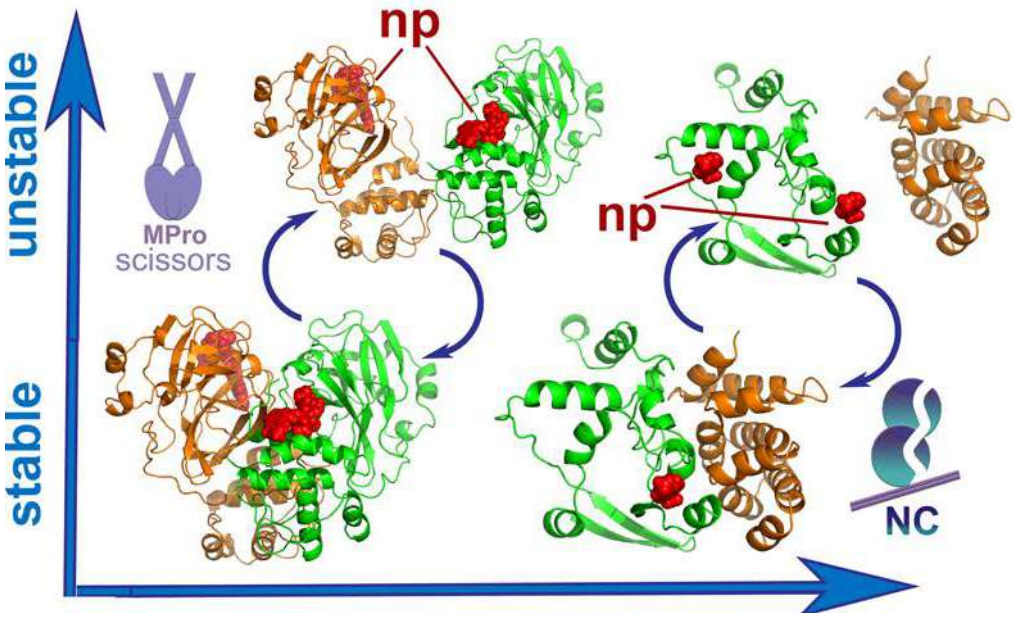
**Nucleocapsid**

**packs and imports the SARS-CoV-2 viral RNA into the host cell nucleus (dimer, binds to importins)**

# The targets


SARS-CoV-2
viral proteins
viral RNA
MPro scissors
Spike
membrane
envelope
RNA
nucleocapsid
NC
translation
production
$-\frac{dV}{dr} = m\frac{d^2r}{dt^2}$
dynamics
NC
viral proteins
MPro scissors
01 00 11 10
computer digits
HPC
HPC
HPC

# The targets




**natural products (anti-viral drugs)**

**SARS-CoV-2**

https://arxiv.org/abs/2007.07736

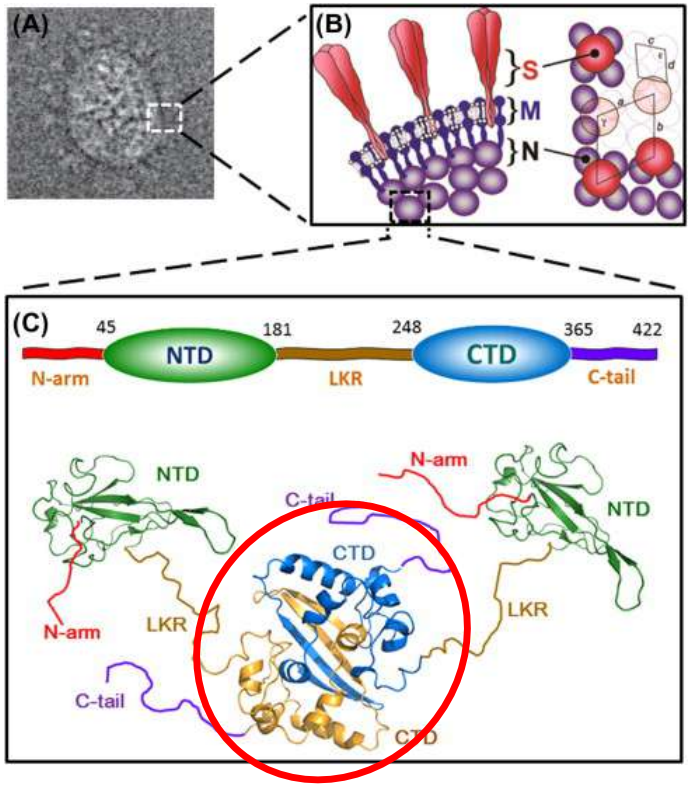
➔ **guided experimental studies**

**dimerization of the SARS-CoV-2 main protease**

**Interaction of nucleocapsid protein with importins that pack and transport the viral RNA into the host cell nucleus**

# The targets / Docking

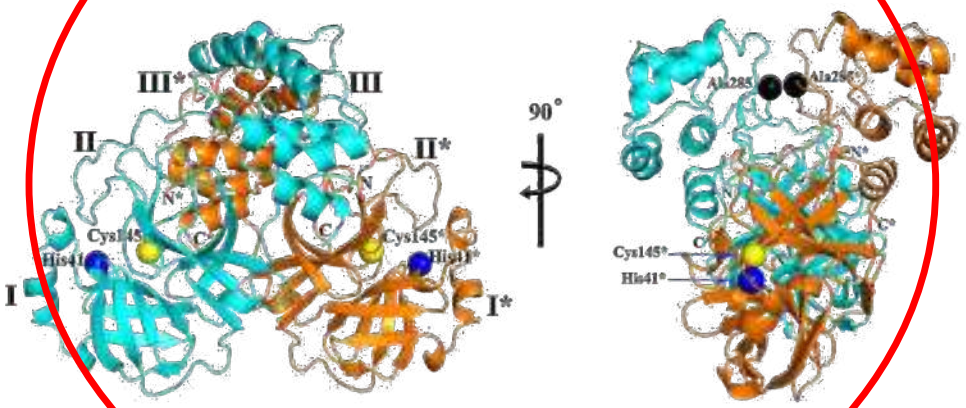
## N – Nucleocapsid protein
**binding of p-cymene**




doi: 10.1016/j.antiviral.2013.12.009

NC/ p-cymene: 277, 278, 280, 283, 284, 285, 310, 313, 330, 331, 333, 340, 356, 362, 366, 378, 380

## 3CL-PRO –3C-like proteinase
**binding of Fortunellin**
**(a natural polyphenol)**




doi: 10.1126/science.abb3405

Mpro-Fortunellin : 44, 48, 53, 82, 83, 84, 111, 112, 118, 137, **138, 139, 141**, 159, 182, 238, 239, 240, 286, 287, 288, **289, 291** ➔ No mutations (T45, S46, E47, L50)

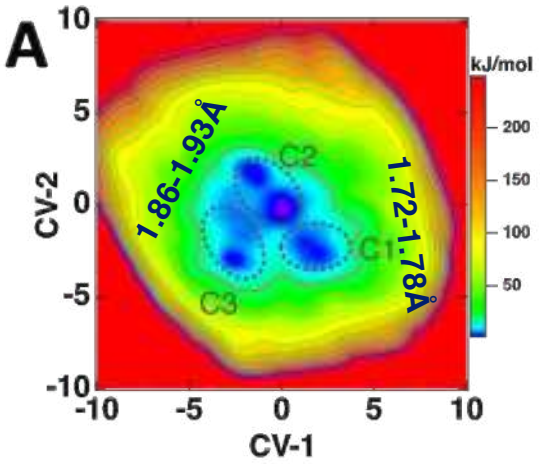
doi:10.3390/ijms21093099

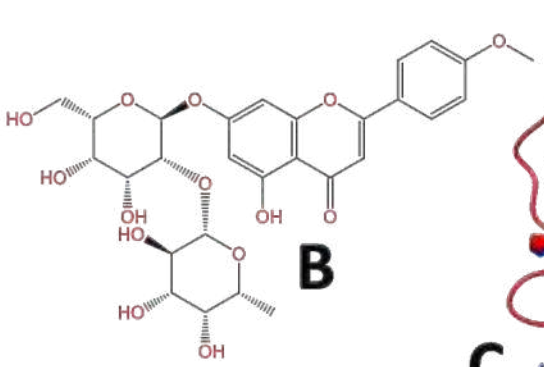
# The main protease

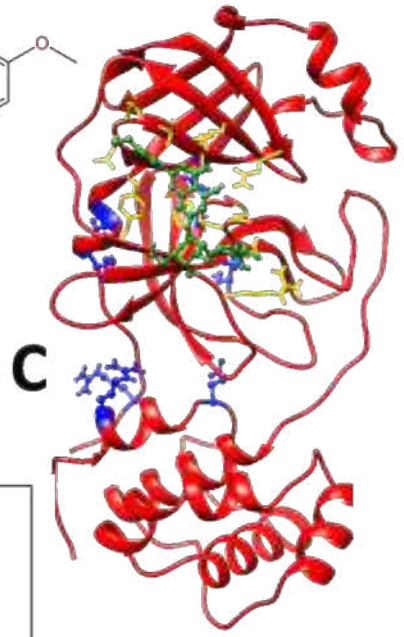
3CL-PRO –3C-like proteinase (90μs)

**binding of Fortunellin weakens dimerization**

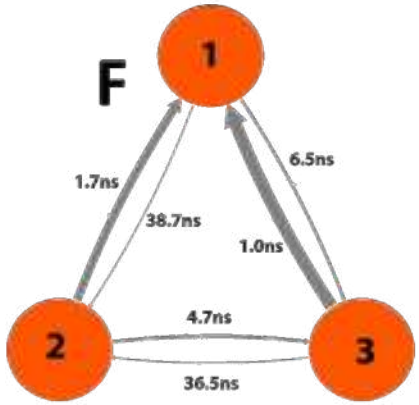
**cross-section (4, 10, 11, 14, 28, 139, 140, 147, 290, 298)**



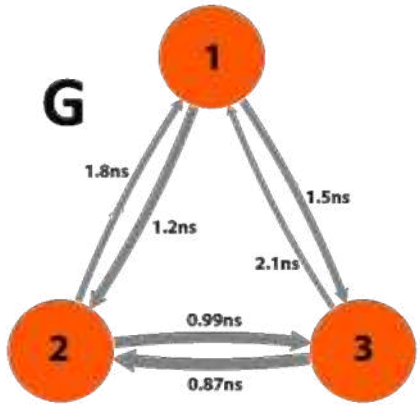



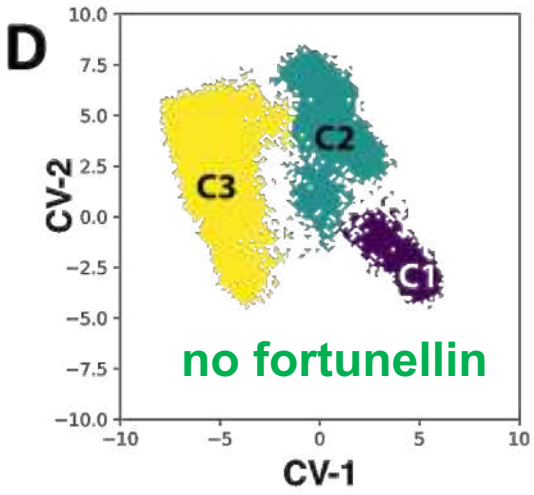



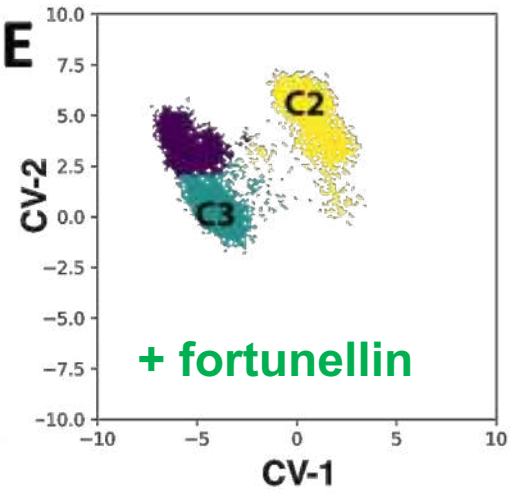

**no fortunellin**




**+ fortunellin**

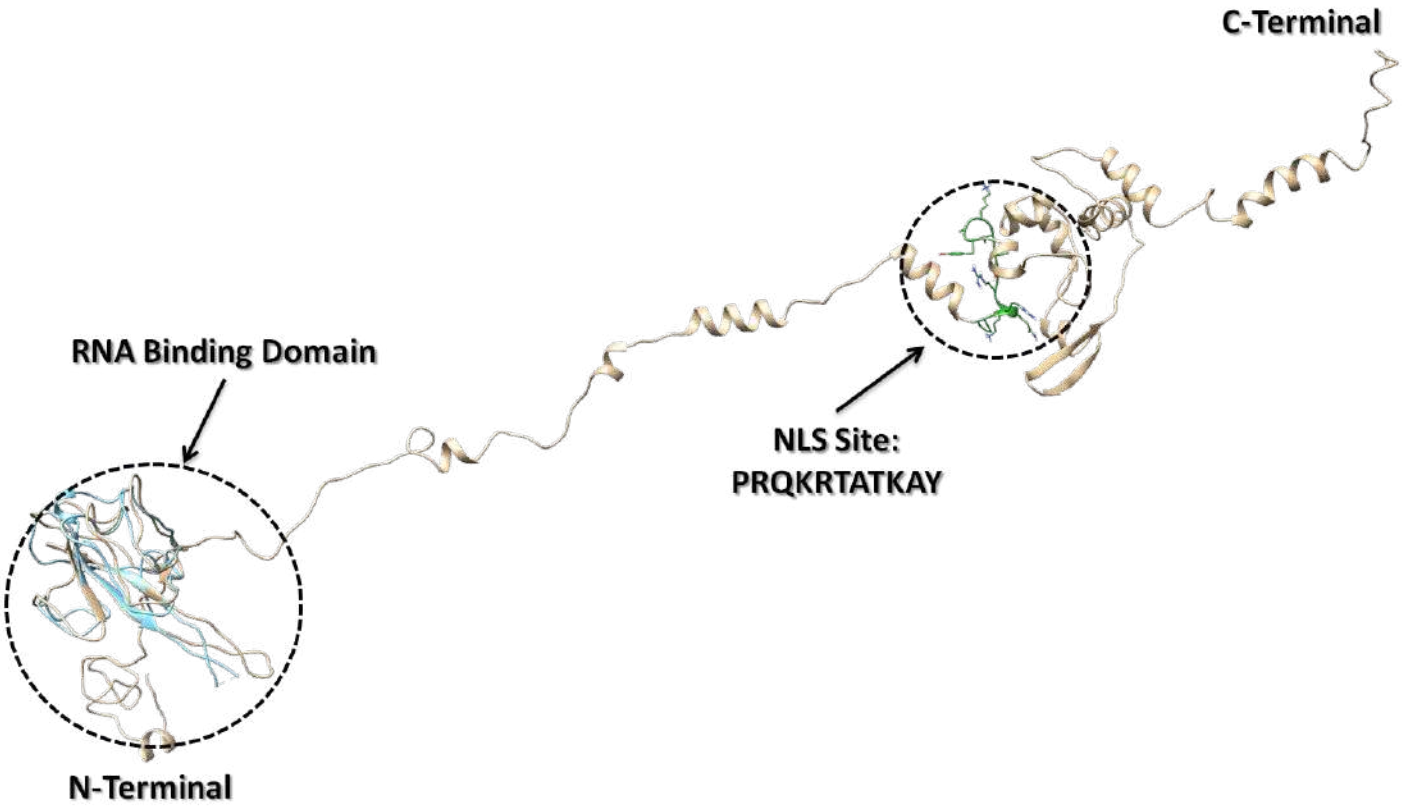


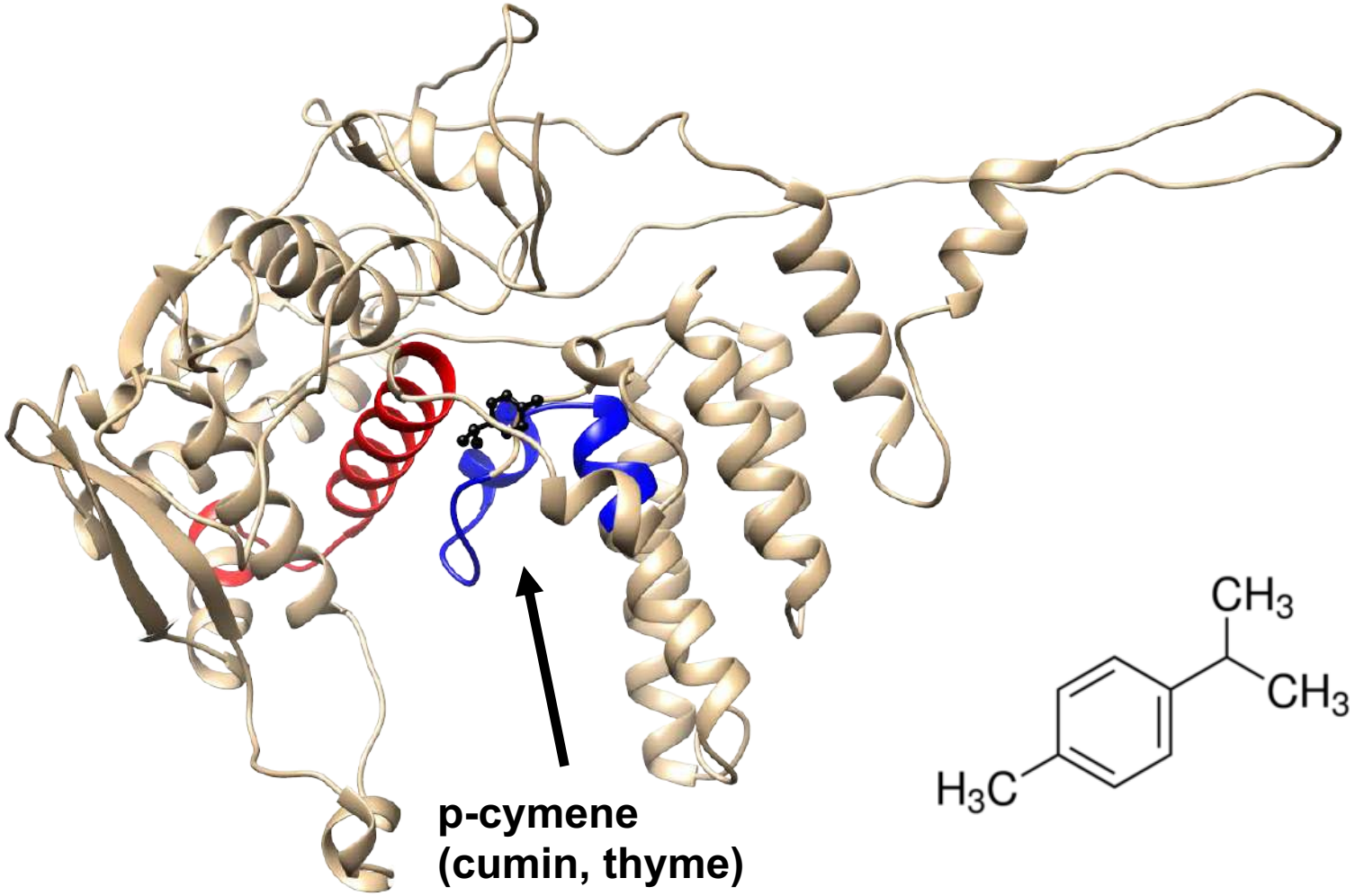




https://arxiv.org/abs/2007.07736

# The nucleocapsid



# The nucleocapsid



# Thank you!


**PROJECT: COVID1941/ RACOV41 – 18.000.000 std cpu core hrs (6 months)**


•**Irene ROME** – ***DARI access in March 2020***

2292 AMD Rome (Epyc) 2.6GHz bi-processor compute nodes, 64 cores/proc -
293,376 compute cores for 11.75 PFlop/s peak power,
•256 GB of DDR4 memory / node,
•InfiniBand HDR100 interconnect.

# Thank you!

PRACE Digest 2020

## Protein targets of SARS-Cov-2

# Using molecular dynamics to find drugs and vaccines for Covid-19

**Molecular dynamics simulations allow us to see into the hidden atomic-scale world that makes up everything we see. Understanding SARS-CoV-2 at this level is helping Vangelis Daskalakis of the Cyprus University of Technology to identify weaknesses in the virus that can be exploited and targeted through drugs and vaccines.**

The Covid-19 pandemic has created a high demand for the characterisation of molecules that can inhibit viral functions, as well as for the production of vaccines that can trigger an efficient and safe immune system response. The rapidly spreading virus has sparked an unprecedented response from the research community as the world faces a health challenge of enormous proportions.

*Vangelis Daskalakis*

Vangelis Daskalakis of the Cyprus University of Technology is a researcher who uses molecular dynamics simulations to study how molecules within biological systems move and interact. These types of simulations can show the role that dynamics play in the function of biological molecules.

Collaborating with colleagues from the University of Crete (School of Medicine, Dept. of Biology), Daskalakis and his team decided to use their skills in the fight against the pandemic. They believe that research on SARS-CoV-2 should focus on proteins

functions. Daskalakis says that any strategy against the disease should include several proteins instead of a single one, as there is a high risk of drug resistance induced by rapid evolution in the viral genetic material. As such, the team has been looking at these two protein targets on the SARS-CoV-2 virus as part of a PRACE project. The main protease is a non-structural protein that cleaves the viral polyprotein into functional proteins – a critical step during viral replication. The second target, the nucleocapsid, is used in the packaging of the viral genome through protein oligomerisation and in keeping it stable inside the virus.

Using 3D structures of these proteins obtained from Cryogenic Electron Microscopy and X-ray crystallography as a starting point, Daskalakis and his team have employed enhanced sampling methods to try and ascertain the shape that these molecules take in vivo. After that, they screened a large database of natural products against these proteins to see if they could find any potential inhibitors to

# Concluding Remarks

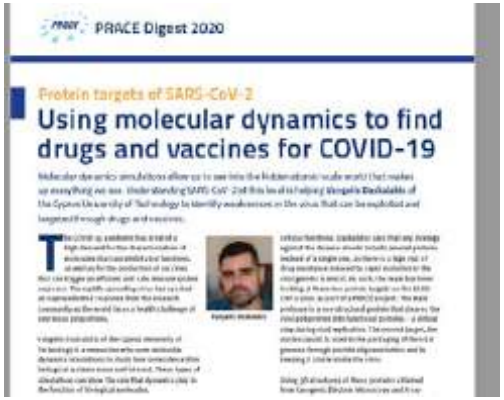
**Molecular Dynamics/ new sampling techniques and High Performance Computing resources: accurate description of the structure and function of large protein complexes at atomic scale, contributing essential knowledge in structural biology**

**Simulation outputs at atomic scale exert consistency with and can guide experiments with considerably lower costs!**

# Research Funding– Thank you!

***2010-2018*** 15.000.000 std-cpu core hrs

***2018-2020*** **~48.000.000 std-cpu core hrs**
**~ €2.700.000 (ACG - acg.maine.edu)**


PRACE Digest 2020

Protein targets of SARS-CoV-2

Using molecular dynamics to find drugs and vaccines for COVID-19


PRACE DIGEST 2017

Let there be light: Photo protection in plants

Plants have a number of adaptations that allow them to thrive in low light conditions, but this means that excess sunlight can be lethal to them. Vangelis Daskalakis of the Cyprus University of Technology has been using molecular dynamics simulations to explore the mechanism via which plants can protect themselves by dissipating excess energy from sunlight as heat.

*"By combining this theoretical computational data with existing experimental data, we are starting to get real insight into the mechanisms of photo protection"*

**POST-DOC/0916/0049**
**€160.000**
**Technology/ΘΕΠΙΣ/0609/05**
**€133.181**

**Start-up Grant (CUT): €40.000**

**Networking (CUT): €40.000**

**(human resources, equipment )**

# Thank you!

**Research Team CEN/CUT (Photosynthesis):**

*simulations*

Dr. Taxiarhis Stergiannakos

Mr. Sotirios Papadatos

*experiments*

Dr. Eleni Navakoudis

Mr. Marios Papageorgiou

**Research Collaborators (Cellular Respiration):**

*experiments*

Prof. Costas Varotsis

Assoc. Prof. Eftychia Pinakoulaki

Dr. Costas Koutsoupakis