How to get time-resolution from frozen images: combining cryoEM with correlated scattering

Sebastian Doniach
(Stanford University, USA)

Overall Purpose:

Structure and interactions of Biomolecules

why combine cryoEM with xFEL scattering?

multiple conformations combined at time of freezing
vs multiple conformations frozen in
time at moment of injection into the xFEL beam

Experimental program:

Study the relation between structure, dynamics, and function, of molecular chaperones

biological functions of proteins

are governed by their three-dimensional fold.

Protein folding, maintenance of proteome integrity, and protein homeostasis (proteostasis) critically depend on a complex network of molecular chaperones.

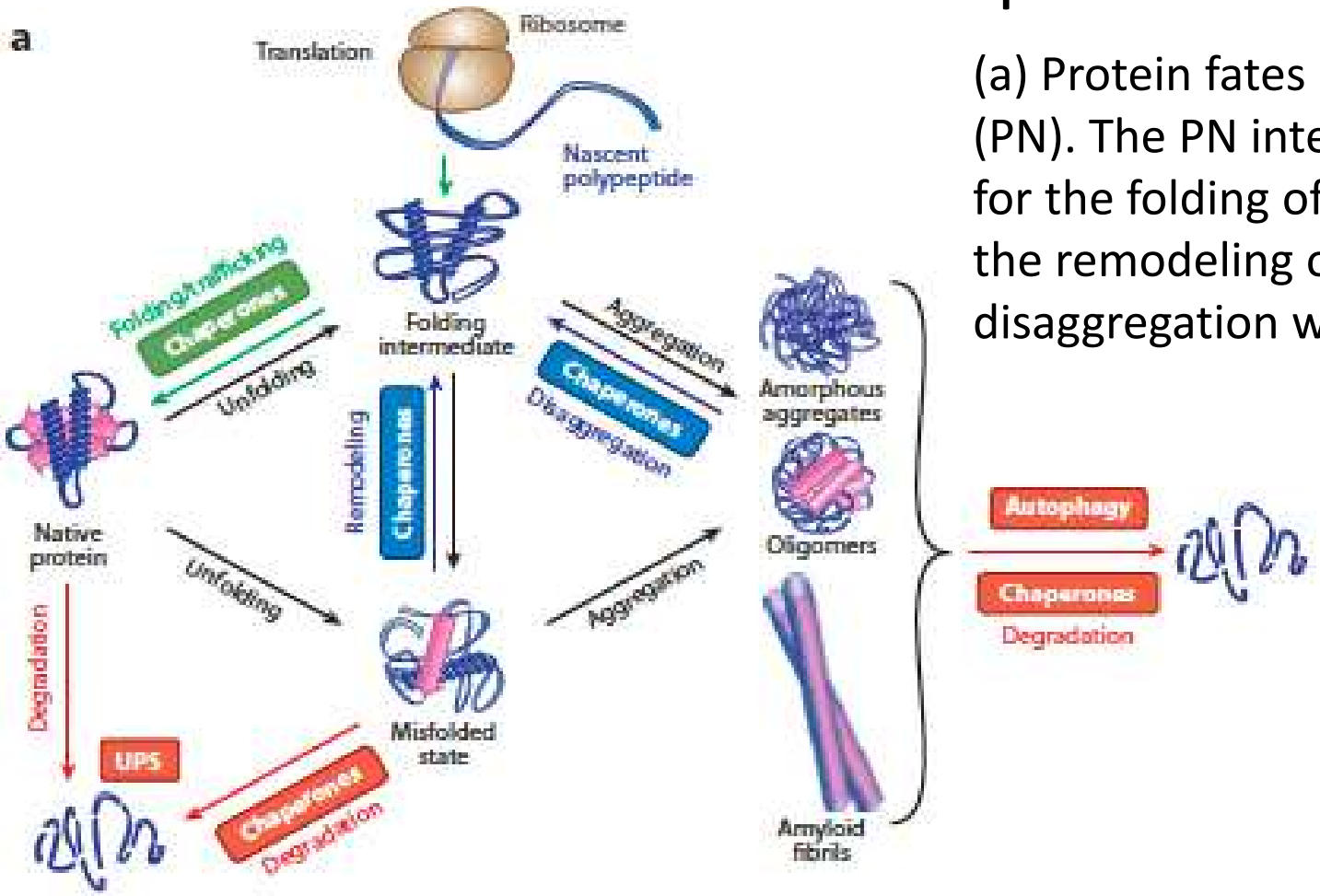
In the cytosol, different classes of molecular chaperones cooperate in evolutionarily conserved folding pathways.

Folding occurs upon controlled release of newly synthesized proteins from these factors or after transfer to downstream chaperones such as the chaperonins

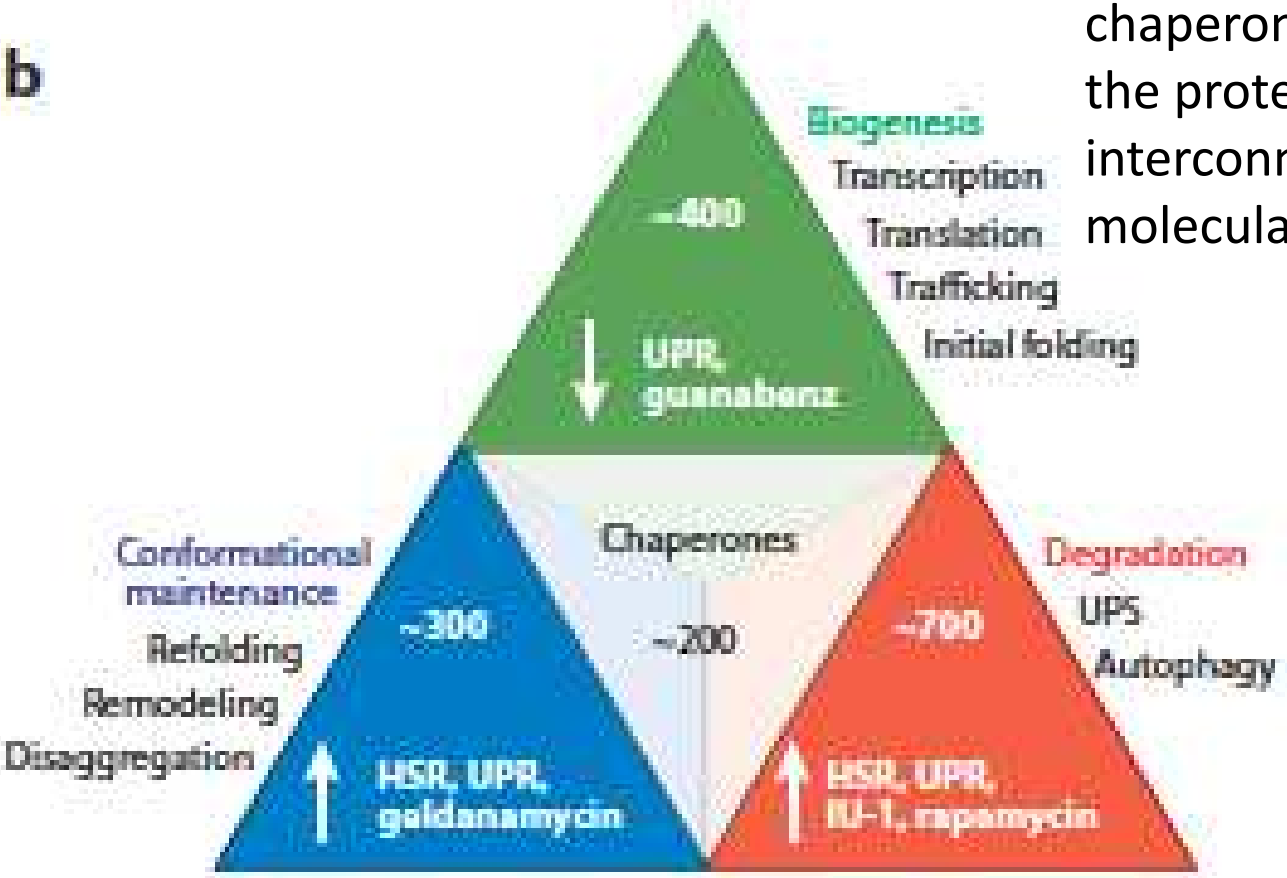
Hsp90: structure and function

Hsp90 is a highly abundant and ubiquitous molecular chaperone which plays an essential role in many cellular processes including cell cycle control, cell survival, hormone and other signalling pathways. It is important for the cell's response to stress and is a key player in maintaining cellular homeostasis.

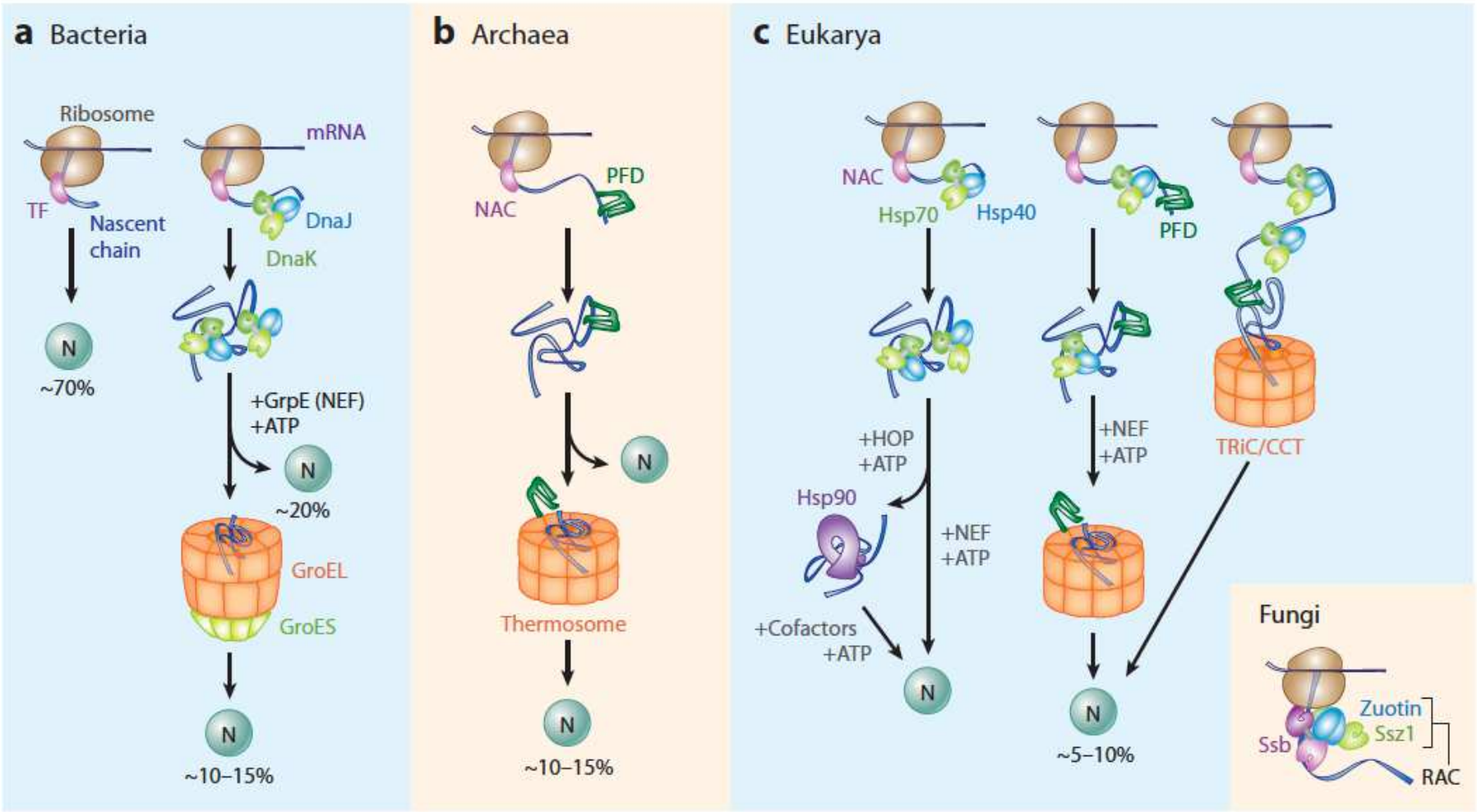
The proteostasis network



(a) Protein fates in the proteostasis network (PN). The PN integrates chaperone pathways for the folding of newly synthesized proteins, the remodeling of misfolded states, and disaggregation with protein degradation.



(b) Central role of molecular chaperones. The three branches of the proteostasis network are interconnected by the functions of molecular chaperones.



Organization of chaperone pathways in the cytosol. In bacteria (*a*), archaea (*b*), and eukarya (*c*), ribosome-bound chaperones [trigger factor (TF) in bacteria, nascent-chain-associated complex (NAC) in archaea and eukarya], aid folding cotranslationally by binding to hydrophobic segments on the emerging nascent chains.

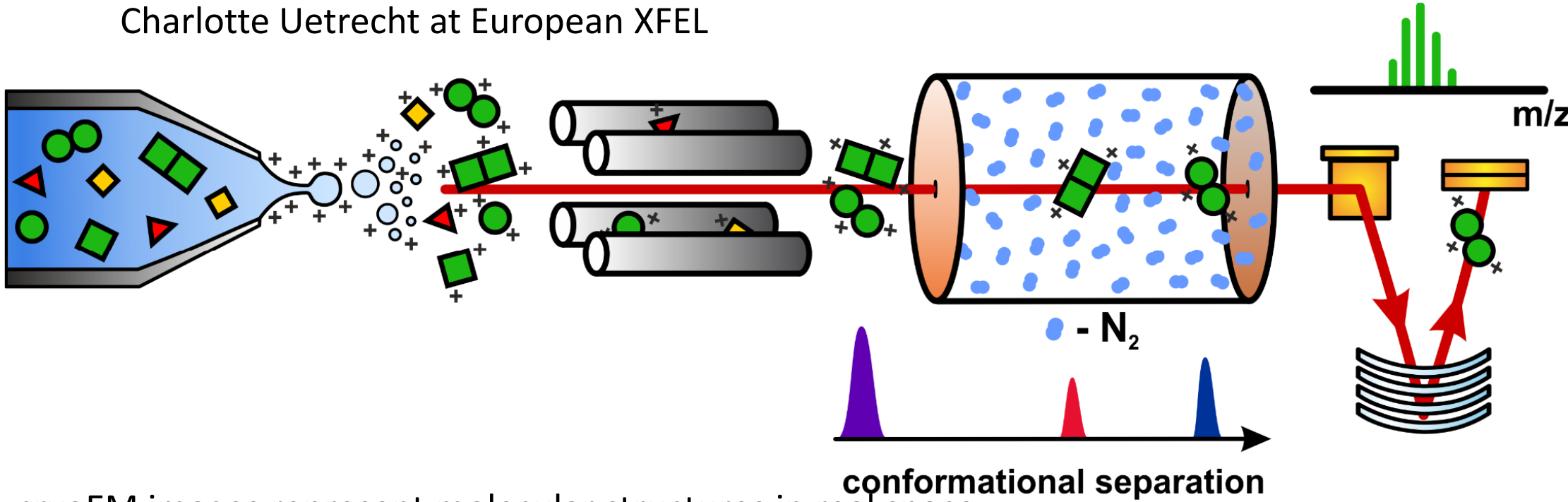
Significance of Conformational Maintenance

After their initial folding and assembly, many proteins remain reliant on molecular chaperones throughout their cellular lifetime to maintain their functionally active conformations.

This is consistent with the notion that proteins with key cellular functions are often structurally dynamic and may be expressed at levels at which they are poorly soluble. Many of the chaperone systems function not only in de novo folding but also in conformational maintenance, i.e., they prevent aggregation of misfolded proteins and mediate their refolding. Specific proteins may interact with as many as 25 different types of chaperones throughout their lifetime, as shown in yeast.

Experimental program (Chuck Yoon group at LCLS)

xFEL scattering electrospray ion trapping – in Collaboration with Charlotte Uetrecht at European XFEL

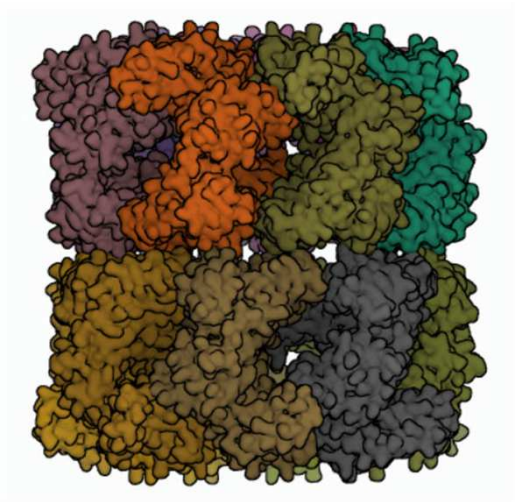


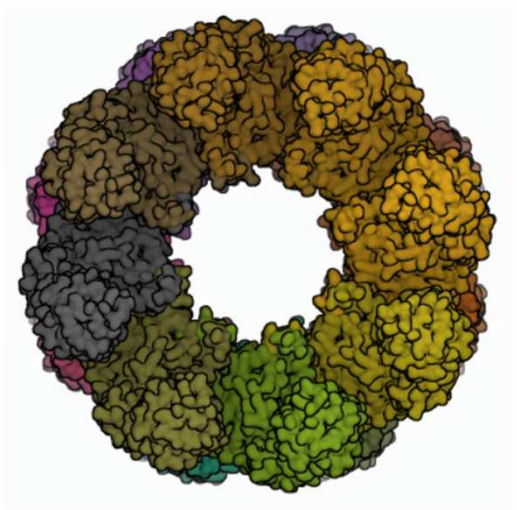
cryoEM images represent molecular structures in real space

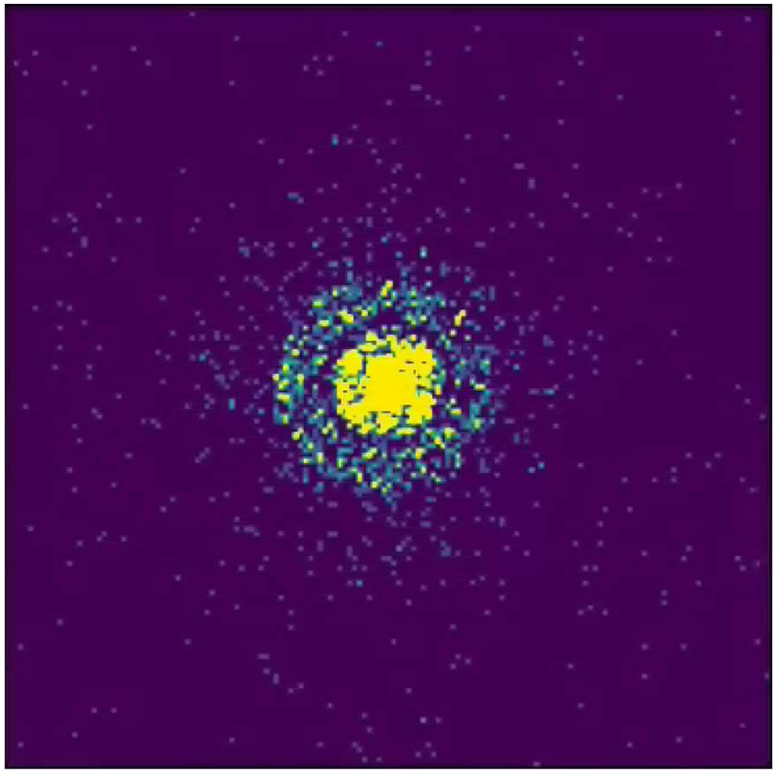
xFEL images represent the sames structures in reciprocal space

a payoff from combining the two imaging methods: a solution to the phase problem.

xFEL scattering simulation – in Collaboration with
Chuck Yoon at SLAC National Accelerator Laboratory







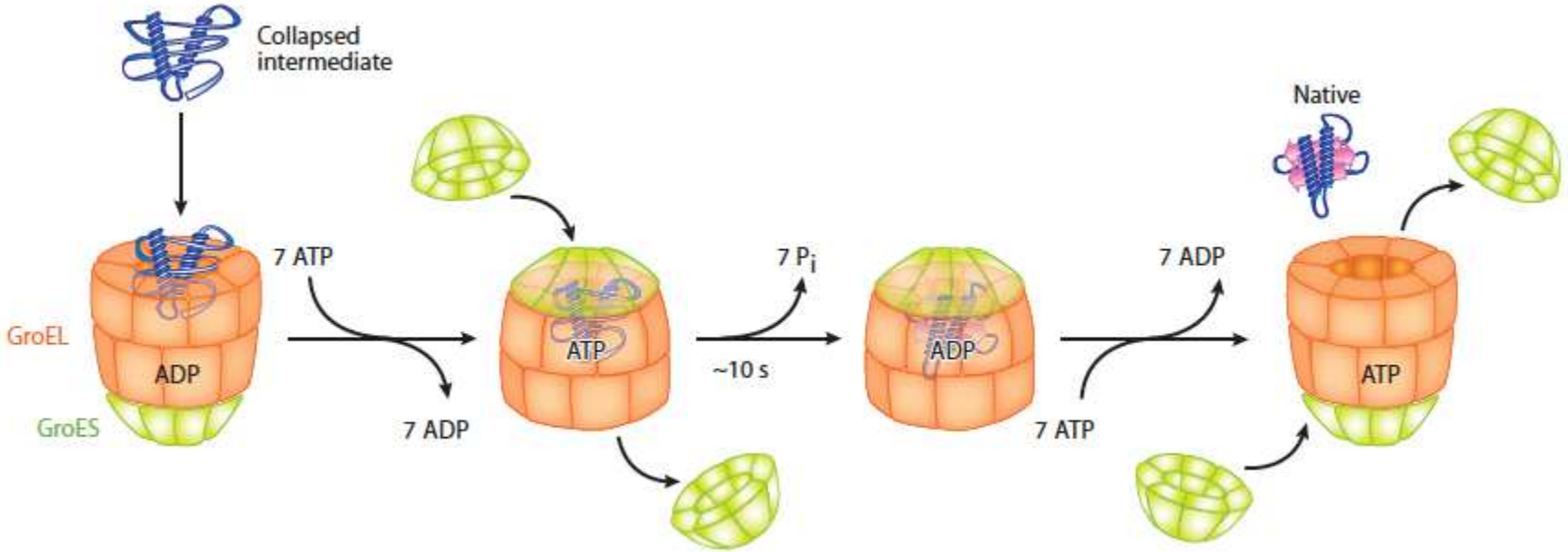
Simulated diffraction pattern from 20 randomly oriented mmCpn particles in the beam using Skopi
https://pypi.org/project/skopi/

Group I cryoEM structuresGroEL-GroES reaction cycle.

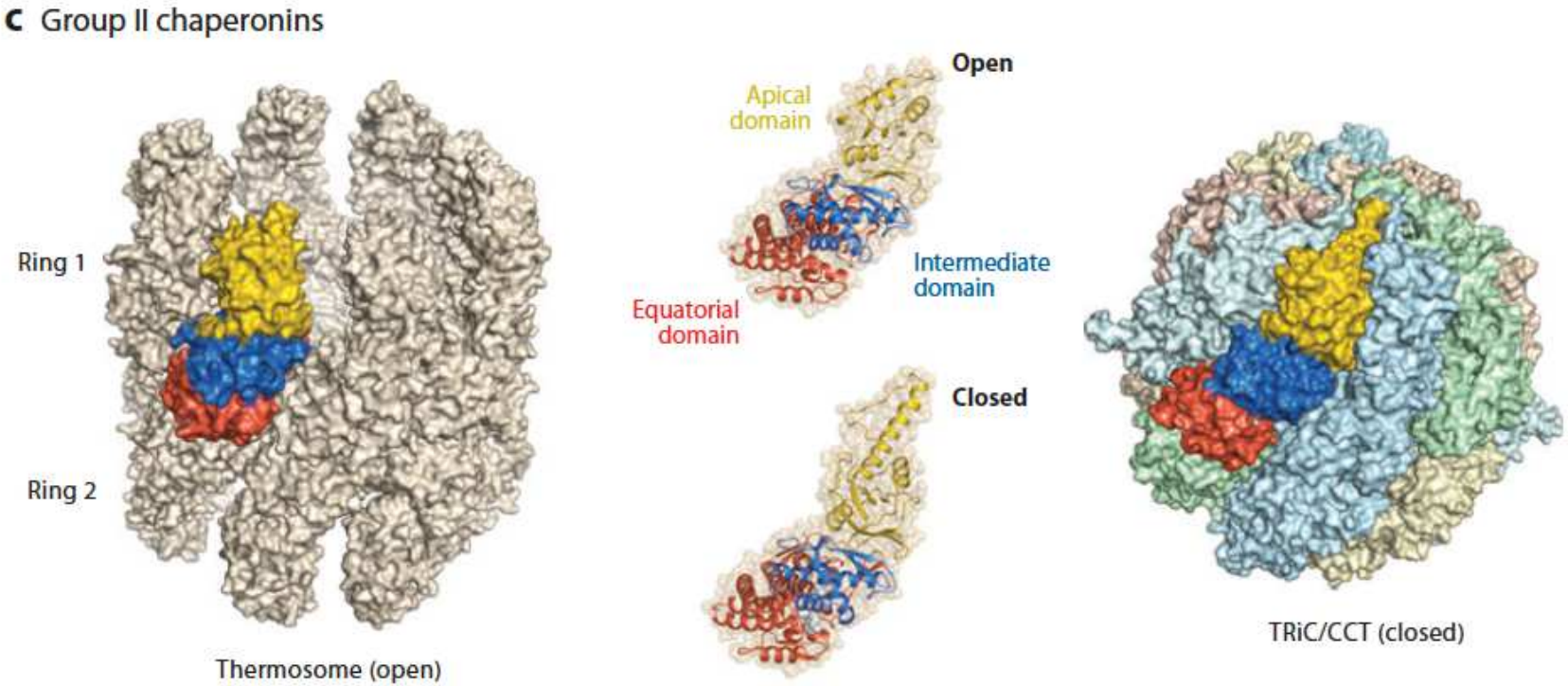
aSubstrate protein as a collapsed folding intermediate is bound by the open GroEL ring of the asymmetrical GroEL-GroES complex, shown in panel a. Binding of ATP to each of the seven GroEL subunits causes a conformational change in the apical domains, which results in the exposure of the GroES binding residues, allowing substrate encapsulation in the cis complex. ADP and GroES dissociate from the opposite ring (trans ring) together with the previously bound substrate.

The newly encapsulated substrate is free to fold in the GroEL cavity during the time needed to hydrolyze the seven ATP molecules bound to the cis ring (10 s). ATP binding followed by GroES binding to the trans ring triggers GroES

bdissociation from the cis ring, releasing the substrate protein



Group II cyoEM structures: MegaDalton scale
(similar to some viruses)

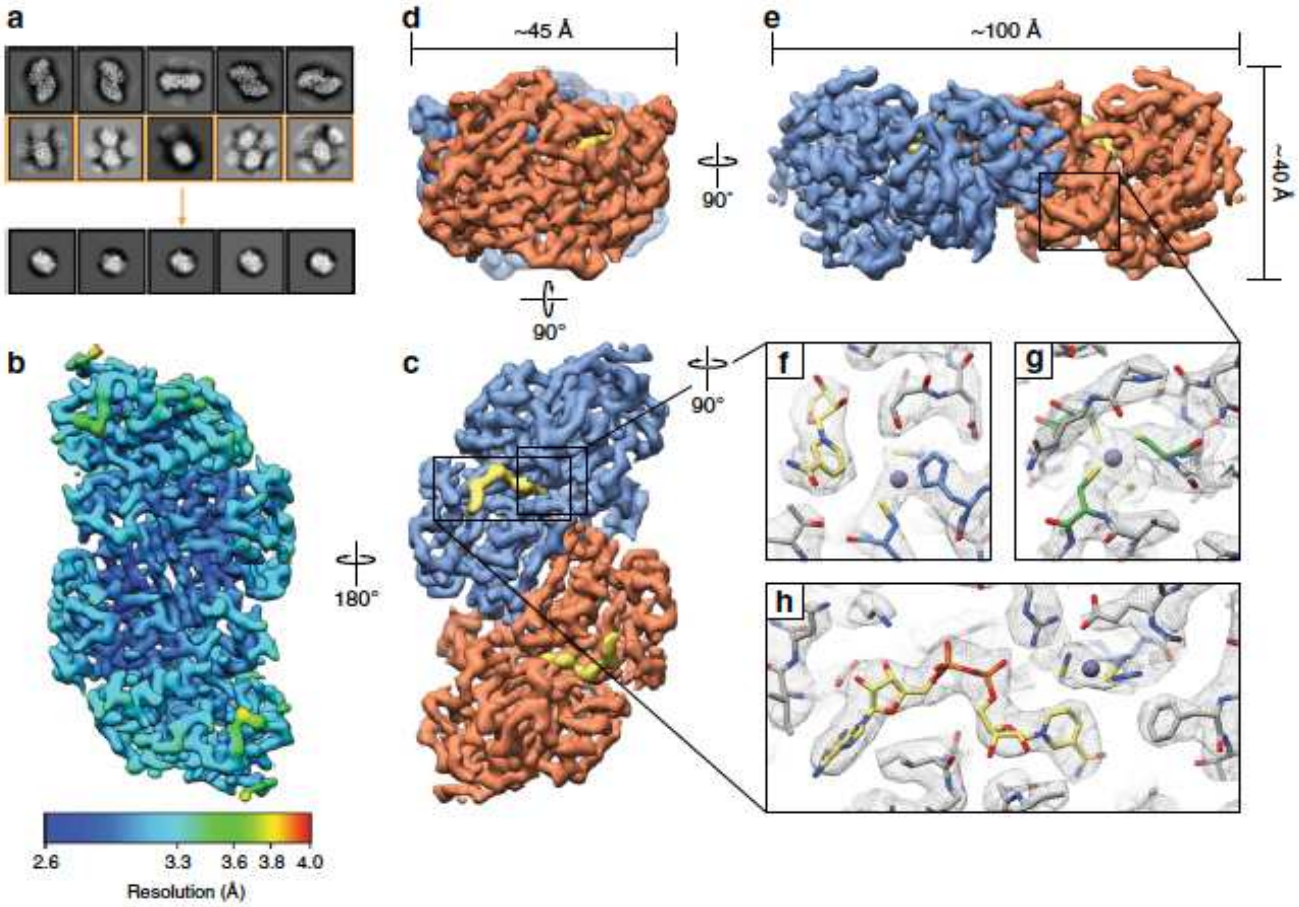


Yujin E. Kim, Mark S. Hipp, Andreas Bracher, Manajit Hayer-Hartl, and F. Ulrich Hartl
Annu. Rev. Biochem. 2013. 82:323–55

Hsp90: (90kD)

The ability to obtain structural data from cryoEM for 80kD molecules enables combining cryoEM data with xFEL scattering data for chaperone proteins in the Hsp90 class.

High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM
Mark A. Herzik Jr., Mengyu Wu1 & Gabriel C. Lander
Nature Communications (2019) 10:1032



CroEM for partlcles smaller than 80kD

2.7 Å resolution cryo-EM reconstruction of 82 kDa horse liver alcohol dehydrogenase. a Representative reference-free 2D class averages of horse liver alcohol dehydrogenase (ADH).

Particles comprising the 2D classes highlighted in yellow were subsequently further classified using a smaller soft circular mask (see methods).

Final cryo-EM reconstruction colored by estimated local resolution estimated with BSOFT[23]

Hsp90: function

The mammalian HSP90 family of proteins is a cluster of highly conserved molecules that are involved in myriad cellular processes.

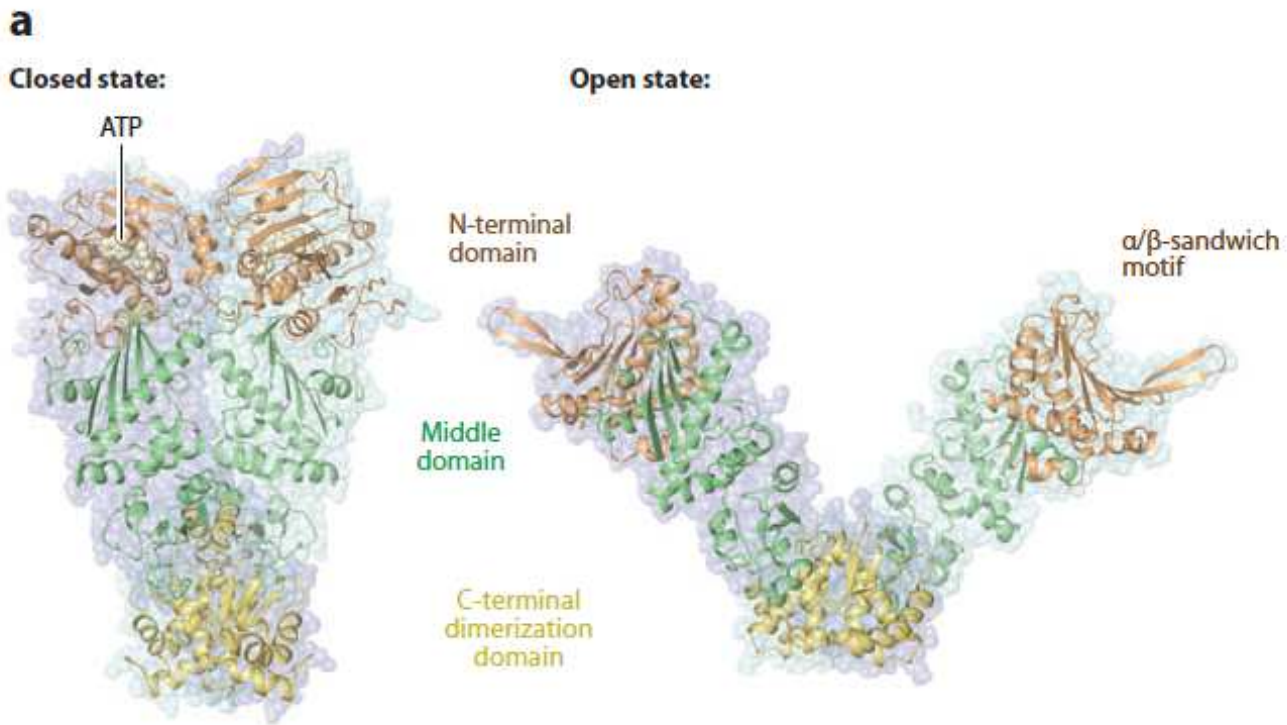
Their distribution in various cellular compartments underlines their essential roles in cellular homeostasis.

HSP90 and its co-chaperones orchestrate crucial physiological processes such as cell survival, cell cycle control, hormone signaling, and apoptosis.

Conversely, HSP90, and its secreted forms, contribute to the development and progress of serious pathologies, including cancer and neurodegenerative diseases.

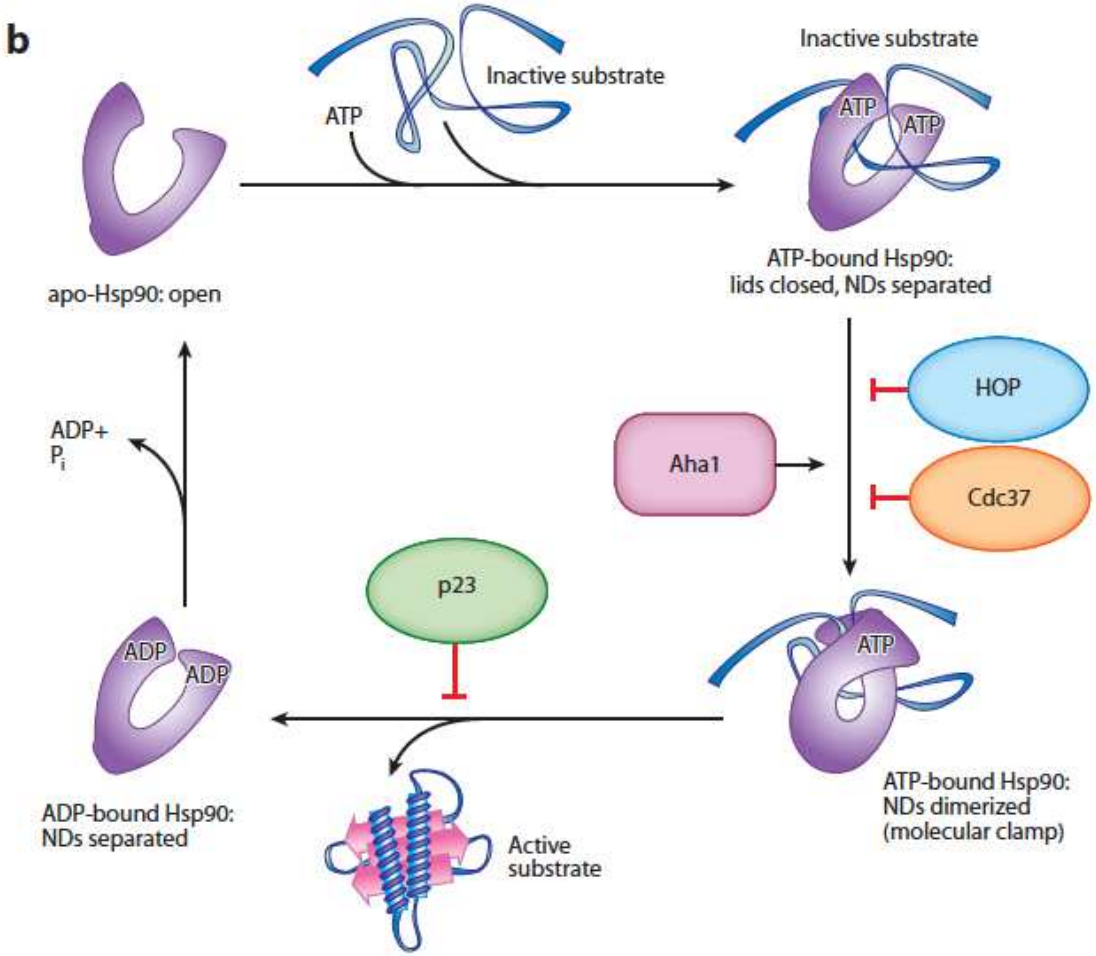
Targeting HSP90 is an attractive strategy for the treatment of neoplasms and other diseases.

Structure and functional cycle of the Hsp90 system. (90 kD)



(a) Structure of Hsp90. Crystal structures of the Hsp90 dimer in the ATP-bound closed state (Saccharomyces cerevisiae; PDB 2CG9) (left) and the nucleotide-free open state (E. coli; PDB 2IOQ) (right) are shown with the nucleotide-binding N-terminal domain in orange, middle domain in green, and C-terminal domain in yellow.

(b) Hsp90 reaction cycle.



Inactive substrate protein binds to ATP-bound Hsp90. In this state the ATP lids are closed and the N-terminal domains are separated.
In the next step, the N-terminal domains dimerize, forming the closed Hsp90 dimer (referred to as a molecular clamp) with twisted subunits.

This metastable state is committed to ATP hydrolysis, upon which the N-terminal domains dissociate. The bound substrate protein is conformationally activated as Hsp90 proceeds through the cycle. Cofactors such as Cdc37 and the Hsp90 organizing protein (HOP) slow the ATP hydrolysis step of the cycle, whereas the activator of Hsp90 ATPase (Aha1) enhances ATP hydrolysis. The cofactor p23 stabilizes the closed dimer to slow the release of substrate protein from Hsp90

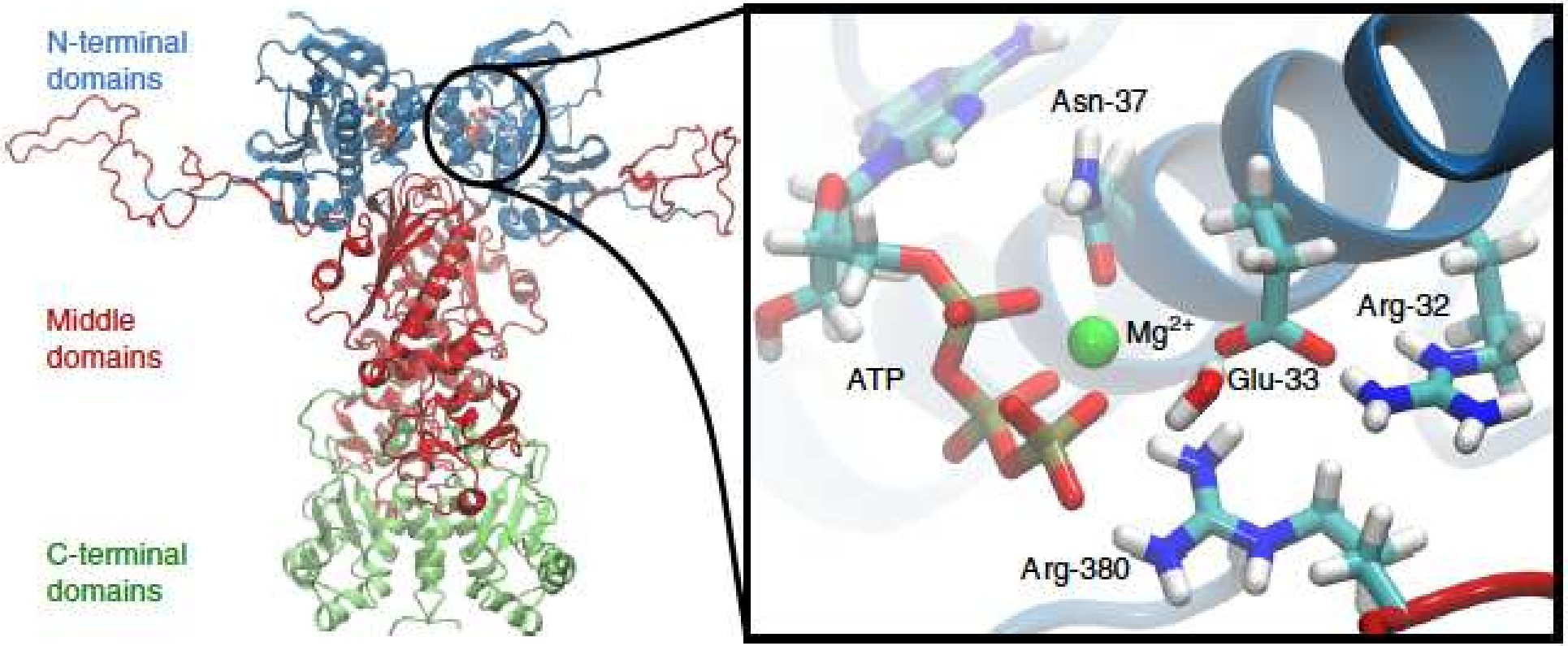


Fig. 1 Structure of the yeast Hsp90 dimer (PDB ID: 2CG9)9. The N-terminal domains of Hsp90 are shown in blue, the middle domains in red, and the C- terminal domains in green. The inset shows a structure of the active site with a bound ATP molecule obtained from an MD simulation, where Asn-37 undergoes a rotation to form a stronger coordination to the magnesium.

Conformational dynamics modulate the catalytic activity of the molecular chaperone Hsp90
Nature communications (2020) 11:1410

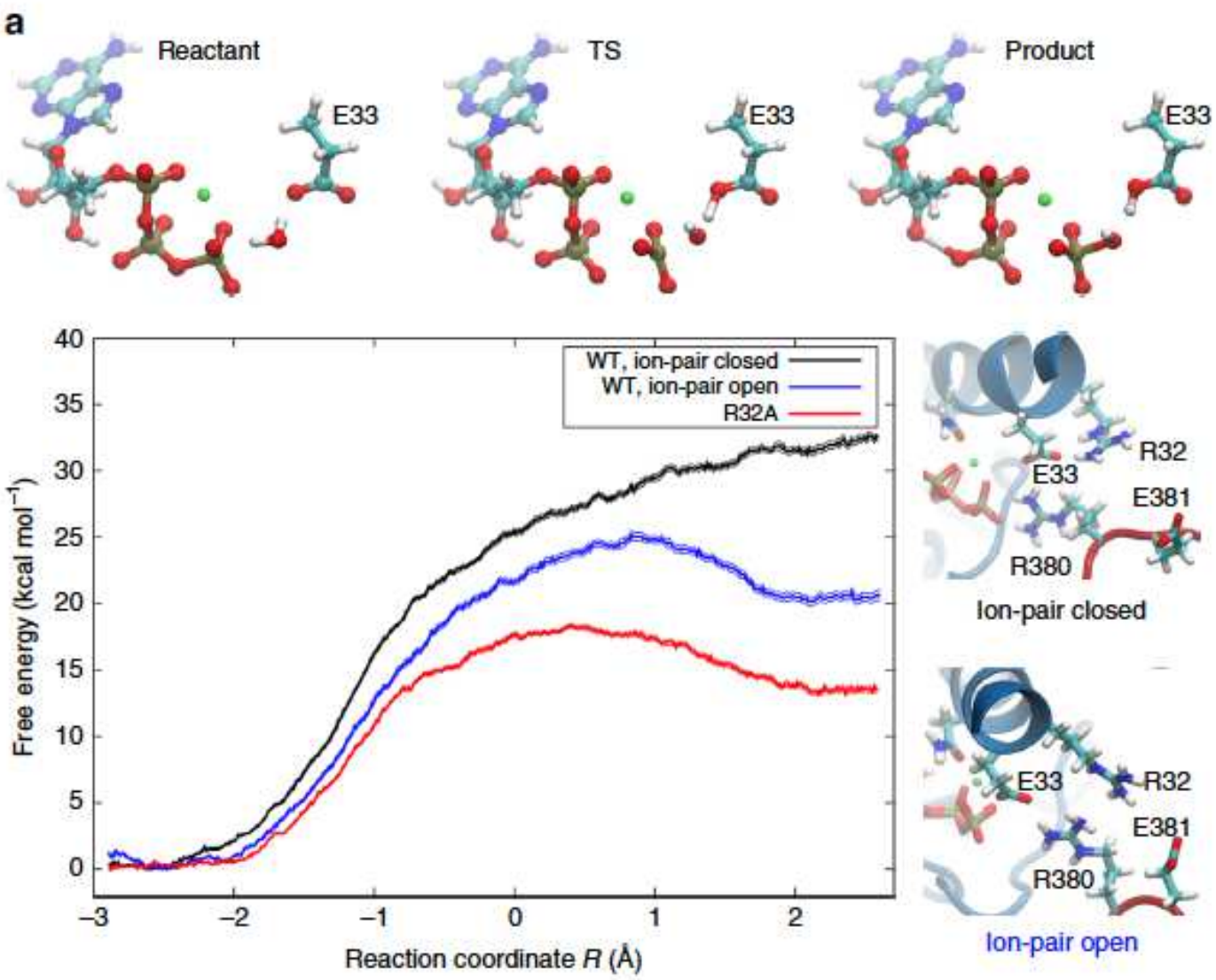
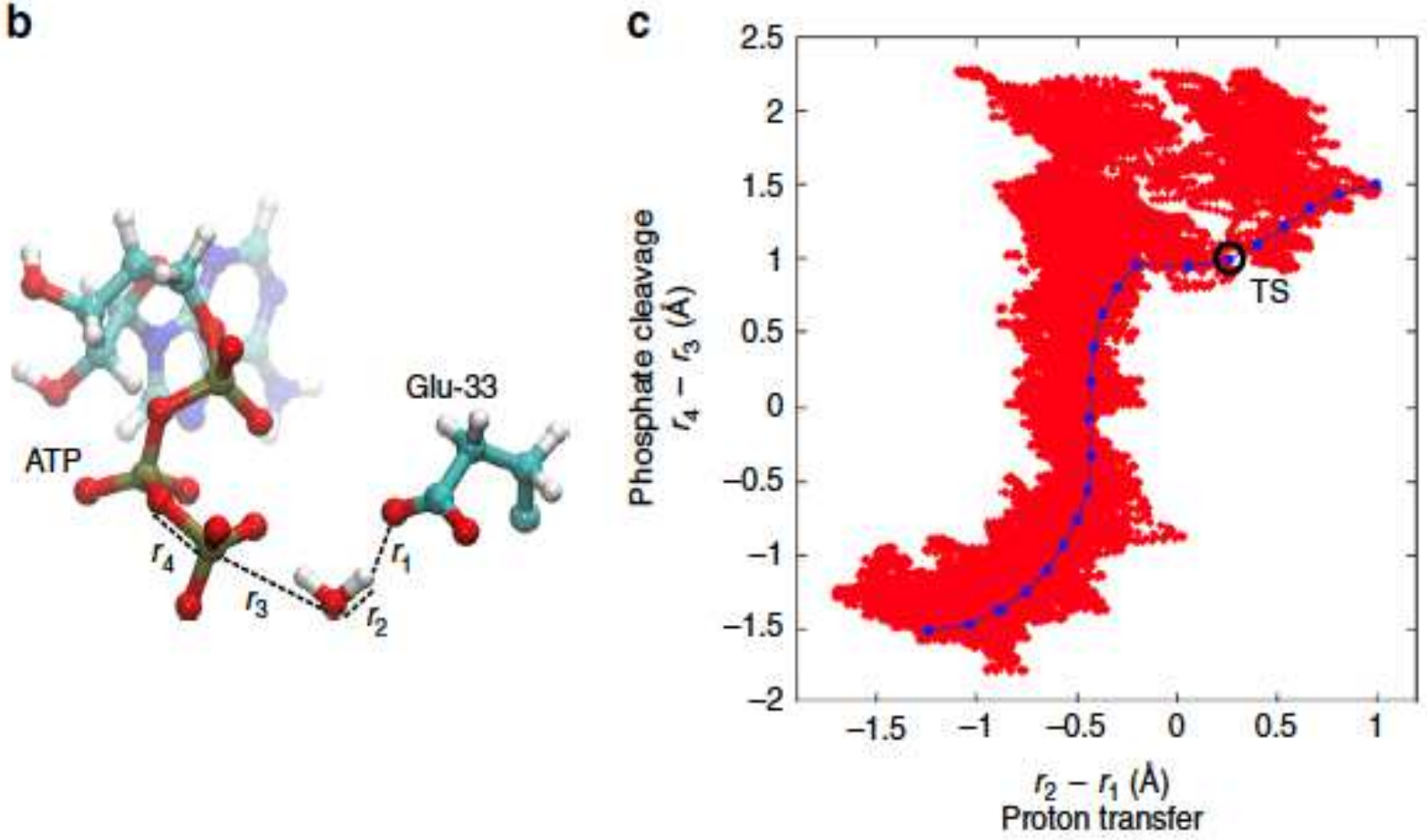


Fig. 2 Energetics and dynamics of ATP hydrolysis in Hsp90. a Reactant, transition state (TS), and product structures showing ATP, Mg2+ (in green), and the sidechain of Glu-33 (E33), extracted from QM/MM calculations of the ATP-hydrolysis reaction.

QM/MM free-energy profiles were calculated with the Arg-32/Glu-33 ion pair closed (black, rR32-E33 < 5 Å) and open (blue, rR32-E33 > 5 Å, see Fig. 3), as well as for the R32A mutant (red).



c Semi-concerted ATP-hydrolysis mechanism from QM/MM free-energy calculations (red dots) and during reaction path optimization (blue dots), showing the sampled reaction coordinates with the Arg-32/Glu-33 ion pair closed. The transition state (TS) of the reaction path optimization is marked with a black circle.

Interactions between biomolecules

Modelling Chaperone machinery allows study of how Proteins bind to one another in specific ways to form protein complexes. The resulting complexes are essential for virtually all cellular processes, and the targeted blocking of protein-protein interactions is a key strategy in modern drug design

Simulation is key to interpretation of XFEL scattering data. Applications to solution scattering data face considerably challenges, in particular for overfitting.

A significant development is in the use of AI (neural network) technologies.

A recently published article is noteworthy: Here a machine learning method is introduced that learns directly from the 3D positions of all atoms to identify accurate models of protein complexes, without using any precomputed physics-inspired or statistical terms. The neural network architecture introduced combines multiple ingredients that together enable end-to-end learning from molecular structures containing tens of thousands of atoms. The architecture is based on equivariance with respect to rotation and translation, local convolutions, and hierarchical subsampling operations.

"Hierarchical, rotation-equivariant neural networks to select structural models of protein complexes"
Stephan Eismann, Raphael J.L. Townshend, Nathaniel Thomas, Milind Jagota, Bowen Jing, Ron O. Dror, *Proteins. 2021;89:493 501*

A couple of applications to cancer:

The search for small molecule drugs to impede metastasis

Hsp90 and cancer (1)

The glendamaycin-derived HSP90 inhibitors have been identified as being synthetically lethal with the transcription factor NRF2. They have been utilized in both monotherapy and combination therapies in clinical trials up to phase III for a range of tumor types, including multiple myeloma, non-small-cell lung carcinoma, acute myeloid leukemia, and gastrointestinal stromal tumors. This makes them ideal candidates for drug repositioning as novel treatments for the orphan KEAP1-NRF2 pathway in cancer.

Aberrant activation of Nrf2 is associated with poor prognosis.

Nrf2 is a transcription factor that stimulates the expression of genes which have antioxidant response element-like sequences in their promoter. Nrf2 is a cellular protector, and this principle applies to both normal cells and malignant cells. While healthy cells are protected from DNA damage induced by reactive oxygen species, malignant cells are defended against chemo- or radiotherapy.

Molecular and Cellular Biology November 2020 Volume 40 Issue 22 e00377-20

Hsp90 and cancer (2)

The transcription factor NRF2 (nuclear factor erythroid 2) is the master regulator of the cellular antioxidant response. Though recognized originally as a target of chemo-preventive compounds that help prevent cancer and other maladies, accumulating evidence has established the NRF2 pathway as a driver of cancer progression, metastasis, and resistance to therapy. Recent studies have identified new functions for NRF2 in the regulation of metabolism and other essential cellular functions, establishing NRF2 as a truly pleiotropic transcription factor.

NRF2 and the Hallmarks of Cancer *Cancer Cell 34, July 9, 2018*

Nrf2 activates several oncogenes unrelated to the antioxidant activity, such as Matrix metallopeptidase 9 (*MMP-9*), B-cell lymphoma 2 (*BCL-2*), B-cell lymphoma-extra large (*BCL-xL*), Tumour Necrosis Factor *α* (*TNF-α*), and Vascular endothelial growth factor A (*VEGF-A*).

Cancers (Basel). 2019 Nov; 11(11): 1755.

summary:

interaction between biomolecules constitutes the machinery of living systems.

function at the cellular level is based on synthesis and interactions of proteins at specific times and places during the cell cycle.

chaperone machinery has evolved to ensure appropriate function of protein machinery and to minimize breakdown of cellular function.

combining the technologies of cryoEM and xFEL solution scattering will increase the knowledge needed for control of cellular function.

acknowledgments:

the work reported here results from a collaboration between the group of Chuck Yoon, and staff at the LCLS, Judith Frydman at the Stanford Biology Department and Yiorgo Skiniotis at the Stanford Medical School.

Our group is benefiting from a collaboration with the group of Charlotte Uetrecht at the European XFEL.