Challenges for molecular structure determination by single particle cryo-EM

Yifan Cheng

Department of Biochemistry & Biophysics
University of California San Francisco

NRAMM cryo-EM workshop
November 9-14, 2014

A list of topics to talk about

5:00 pm **Challenges for Molecular Structure Determination** **Yifan Cheng**

[This introductory talk will describe the big picture: what have we achieved and what do we still need to do. Since the last workshop we have seen major advances in image acquisition resulting in (almost?) atomic level structure determination of a number of specimens. Some of the specimens were well behaved whereas others required considerable biochemical characterization and modification before they were suitable for high resolution imaging and analysis. What limited the resolution in the first place with these "difficult" specimens? Was there a strategy for making them better or was it hit and miss? From this work have we learned any general principles that might be applied to other "difficult" specimens? Were there any special approaches involved in imaging or data processing? What are the prospects for a general high resolution methodology for, e.g. single particles? What are the current limitations? Perhaps illustrate these points with great successes and embarrassing failures (without offending people). Do the problems get bigger as we approach atomic resolution? What are the challenges in looking at intermediate states? Where should we focus our efforts in the immediate future? What should we pay attention to and what should we NOT do?]

A few highlights of the last workshop (two years ago)

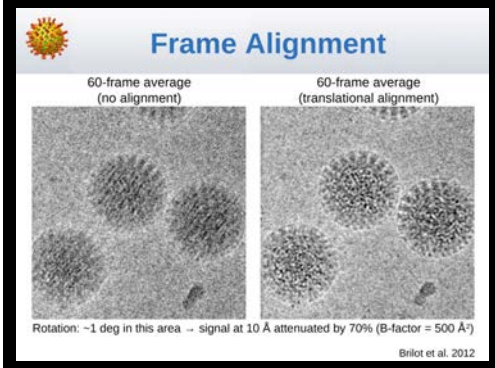
What were presented in the last workshop two years ago (in addition to new approaches and optimization of single particle cryo-EM, from sample preparation, image acquisition and processing, to validations):

* using direct electron detection cameras for single particle cryo-EM (Niko);
* motion corrections plays significant role in achieving high resolution (Niko, Yifan);
* 3D classification (release of RELION, Sjors);

A few highlights of the last workshop (two years ago)

What were presented in the last workshop two years ago (in addition to new approaches and optimization of single particle cryo-EM, from sample preparation, image acquisition and processing, to validations):

* using direct electron detection cameras for single particle cryo-EM (Niko);
* motion corrections plays significant role in achieving high resolution (Niko, Sjors, Yifan);
* 3D classification (release of RELION, Sjors);



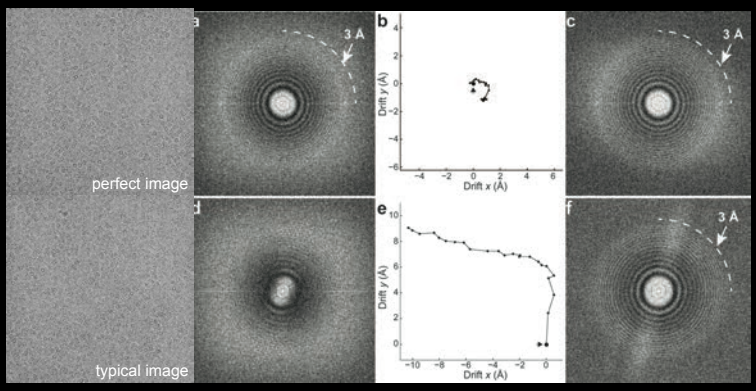
Motion correction was first demonstrated by Grigorieff and Carragher Labs with icosahedral viral particles:

* Brilot et al. "Beam-induced motion of vitrified specimen on holey carbon film" J. Struct. Biol. (2012);

* Campbell et al. "Movies of ice-embedded particles enhances resolution of electron cryo-microscopy", Structure (2012)

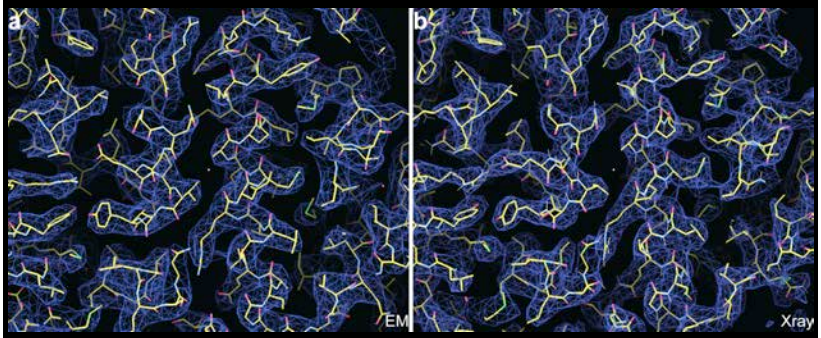
General applications of motion correction

Whole frame motion correction correct for globe and partial local motion, and restore image Thon ring to high-resolution.



Li et al. (2013) Nature Methods

Single particle cryo-EM at crystallographic resolution

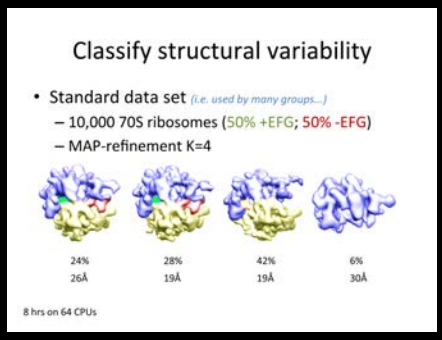


* We determined a 3D reconstruction of archaeal 20S proteasome to the resolution of ~3.3 Å, comparable to the resolution of X-ray crystal structure, 3.4Å.

Li et al. (2013) Nature Methods

Maximum likelihood based classification

Sjors Scheres. “Optimizing image processing”, 2012 NRAMM Workshop.



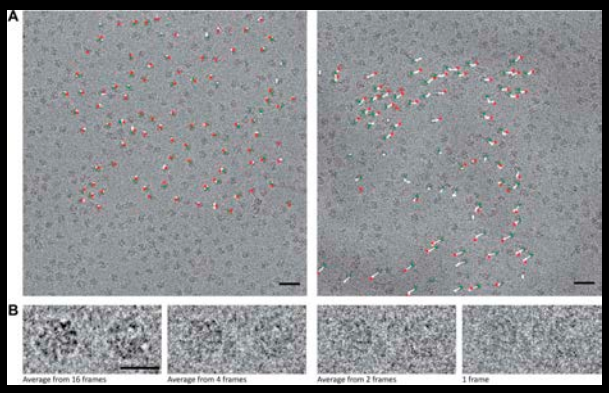
Scheres “RELION: Implementation of a Bayesian approach to cryo-EM structure determination”, J. Struct. Biol. (2012)

Lyumkis et al. “Likelihood-based classification of cryo-EM images using FREALIGN”, J. Struct. Biol. (2013)

General applications of motion correction

Scheres Lab: Bai et al. “Ribosome structures at near-atomic resolution from thirty thousand cryo-EM particles”, eLife (2013)

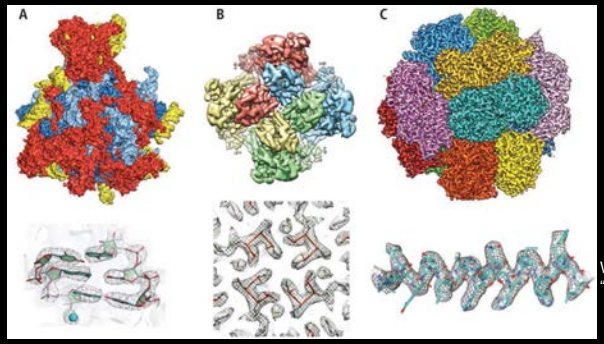
Agard and Cheng labs: Li et al. “Electron counting and beam-induced motion correction enable near-atomic-resolution single particle cryo-EM”, Nature Method (2013)



Bai et al. (2013) eLife

What have “single particle cryo-EM” achieved since then:

- Dose fractionation image acquisition and motion correction become standard procedures.
- Direct detection camera is being used to produce a number of near atomic resolution reconstructions: “Resolution Revolution”



Werner Kuhlbrandt “The Resolution Revolution” Science (2014)

Yeast mitochondrial ribosome, 3.2Å

rat TRPV1 ion channel, 3.4Å

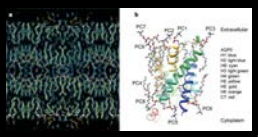
F420-reducing hydrogenase, 3.4Å

Electron crystallography of membrane proteins

Membrane protein structure determination is particularly challenging for X-ray crystallography. It is also challenging for cryo-EM.

* Electron crystallography - 2D and helical crystals;

Bacteriorhodopsin
Light driven H^+ pump

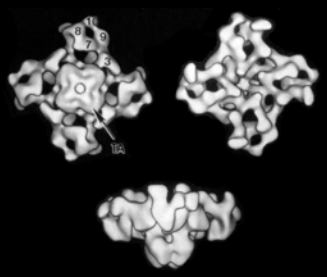


Aquaporin 0
water channel

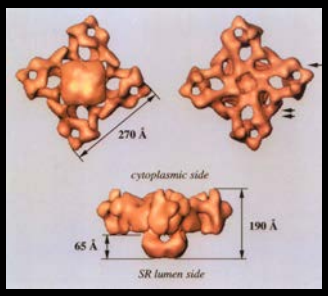
* Crystallographic approach is well-established;
* Most time produced very good structures of membrane proteins at various resolutions.
* Resolution is limited mostly by the quality of crystallinity, and crystallization is still an art.

Single particle cryo-EM of membrane proteins

A long journey that is full of hopes and excitements.



Radermacher et al. (1994)
J. Cell Biol., 127: 411-423

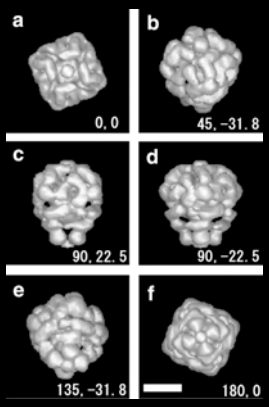


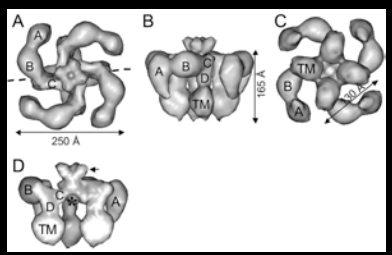
Serysheva et al. (1995)
Nat. Struct. Biol., 2: 18-24

Single particle cryo-EM of ryanodine receptor at ~ 30Å resolution.

Single particle cryo-EM of membrane proteins

But also with embarrassments!



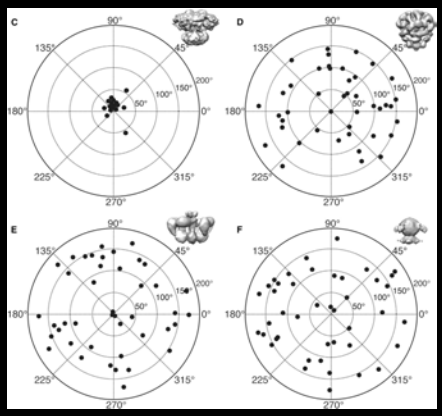


Structure of the type 1 onsoitol 1,4,5-trisphosphate receptor revealed by electron cryomicroscopy. JBC 2003, 278, 21319-22.

Insoitol 1,4,5-trisphosphate receptor contains multiple cavities and L-shaped Ligand-binding domains. JMB 2004, 336, 155-64.

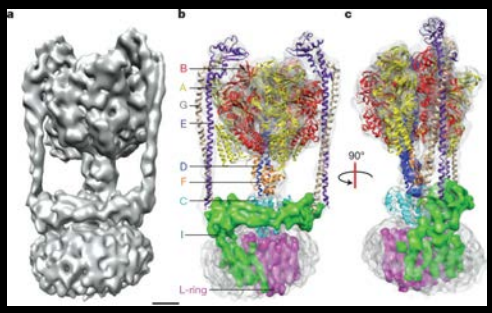
Single particle cryo-EM of membrane proteins

Developments of validation tools!



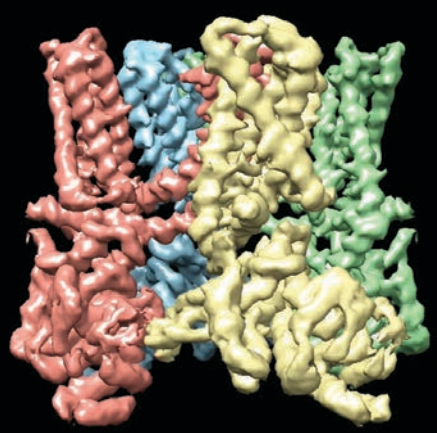
Validation of Cryo-EM structure of IP3R1 channel. Structure 2013, 21, 900-909.

Single particle cryo-EM of membrane protein at sub-nanometer resolution



H^+-driven ATP synthase (9.7Å)
(Lau, et. al. 2011, Nature)

3D reconstruction of TRPV1 ion channel



Raw 3D density map without any modification (resolution filtering and B-factor sharpening). The resolution is estimated at 3.4Å.

Liao, Cai, et al (2013) Nature

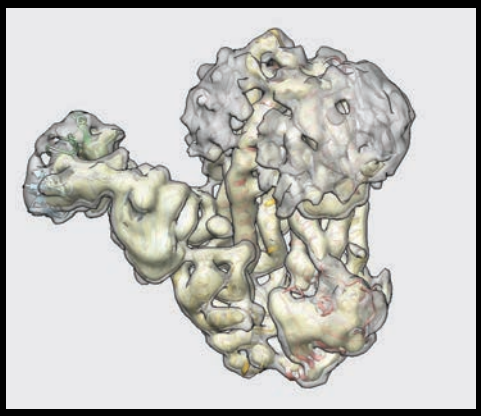
Other recent progresses in membrane proteins

- γ-secretase: Sjors Scheres and Yigong Shi (4 ~ 5Å);
- Mammalian respiratory complex I: Judy Hirst (5Å);
- RyR receptor: Rouslan Efremov and Raunser;
  Joachim Frank and Anrew Marks;
  Sjors Scheres, Yigong Shi and Nieng Yan;

Unpublished, Nieng Yan

Other progresses in membrane proteins

- Glutamate receptor: Sriram Subramaniam and Mark Mayer (7.6 Å);
- ABC exporter: (8.3 Å);



What contributed to TRPV1 structure determination

Contributing factors:

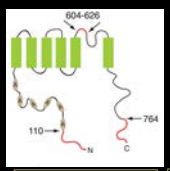
- Production of high quality and biochemically stable proteins;
- Available and well characterized pharmacological reagents;
- Camera related new technologies: high-DQE and dose fractionation;
- Classification of heterogeneous particles;

For ABC exporter reconstruction

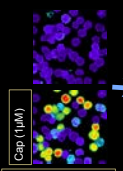
- Fab assisted cryo-EM to study small integral membrane proteins;

Expression and characterization of rat TRPV1

Erhu Cao
David Julius Lab, UCSF



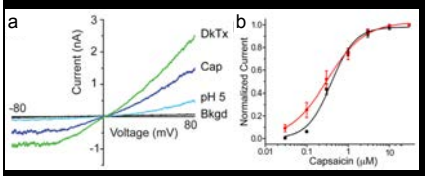
Minimum functional Rat TRPV1 construct



Ca^{2+} image

Expressed and purified from HEK293S

patch clamp of purified TRPV1 reconstituted into liposome

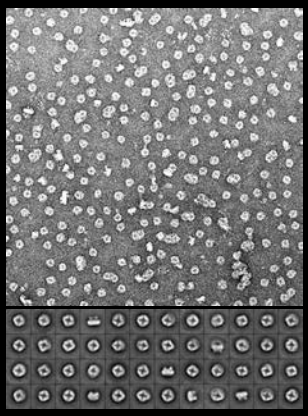


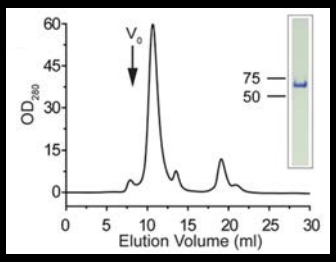
Minimal TRPV1 is fully functional

Dose-responsive curve for minimal and full length (red) TRPV1

Liao, Cao, et al (2013) Nature

Well-behaved TRPV1 proteins

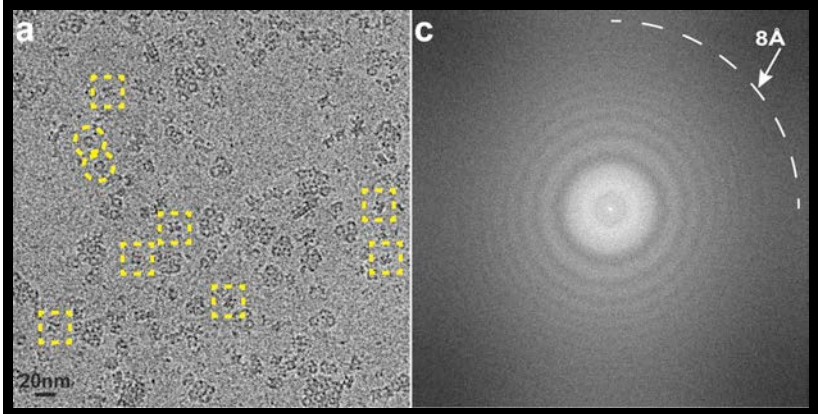




Size exclusion chromatography

Liao, Cao, et al (2013) Nature

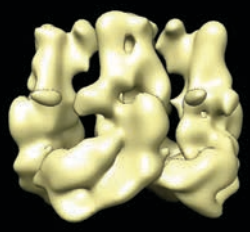
Single particle cryoEM of TRPV1

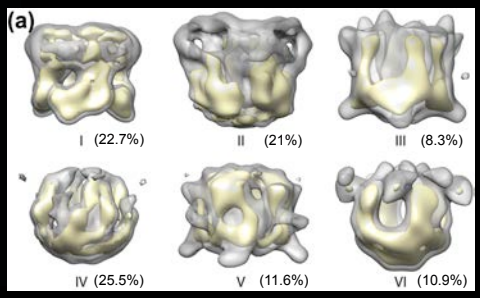


★ Tecnai TF20 microscope operated at 200kV, TVIPS 8K x 8K scintillator based CMOS camera;
★ Image recorded with a defocus of 3.1μm; Thon ring visible at ~8Å resolution;

Liao, Cao, et al (2013) Nature

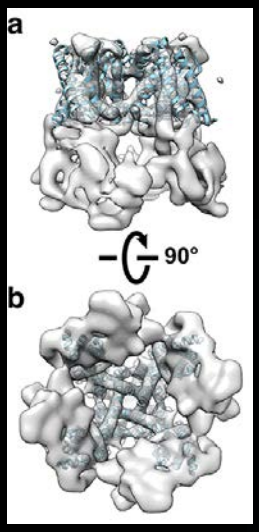
Conformational classification using RELION

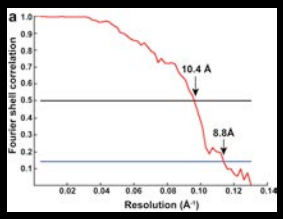


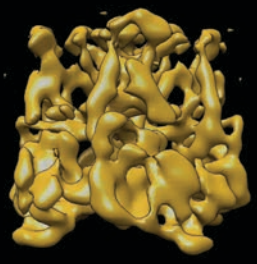


Sjors Scheres: "RELION: Implementation of a Bayesian approach to cryo-EM structure determination", J. Struct. Biol. (2012)

Single particle cryoEM of TRPV1

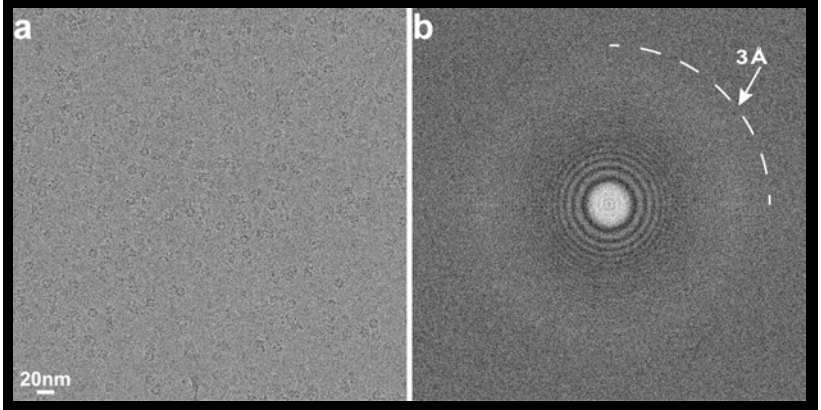






Liao, Cao, et al (2013) Nature

Single particle cryoEM of TRPV1 - using new camera technology



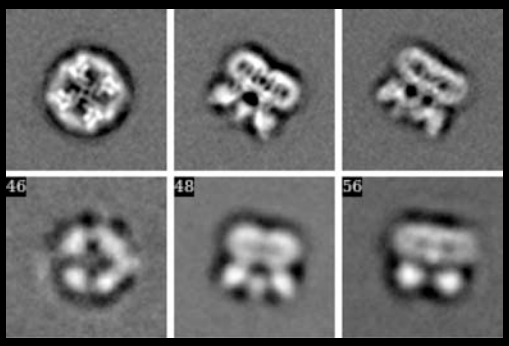
★ Tecnai Polara microscope operated at 300kV, K2 Summit camera;
★ dose fractionation and whole frame motion correction;
★ 6 sec exposure, 0.2 sec frame accumulation (30 subframes per image), total dose: 41$e^-/Å^2$;
★ Image recorded with a defocus of 1.7μm; Thon ring visible at close to 3Å resolution;

Liao, Cao, et al (2013) Nature

Significant improvement of data quality

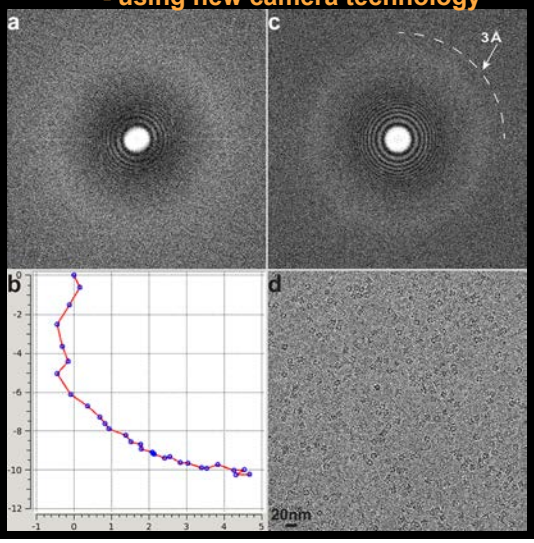
K2 Summit, whole frame motion correction

TVIPS 8K scintillator based CMOS camera



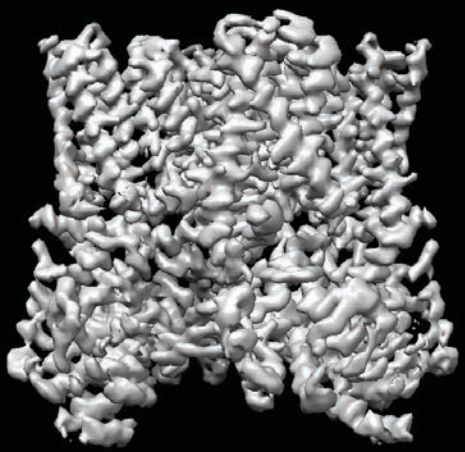
Liao, Cai, et al (2013) Nature

Single particle cryoEM of TRPV1 - using new camera technology



Liao, Cao, et al (2013) Nature

Structure of TRPV1

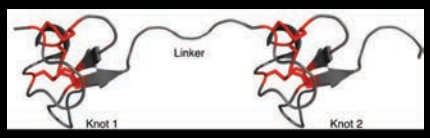


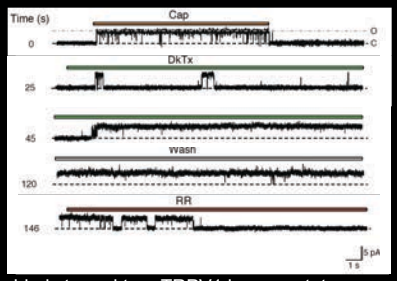
Liao, Cao, et al (2013) Nature

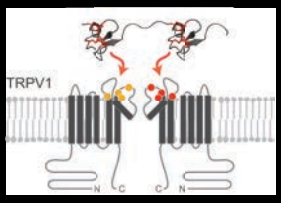
Double Knot Toxin (DkTx)

Ornithoctonus huwena

(Chinese Bird Spider)
(Earth Tiger Tarantula)





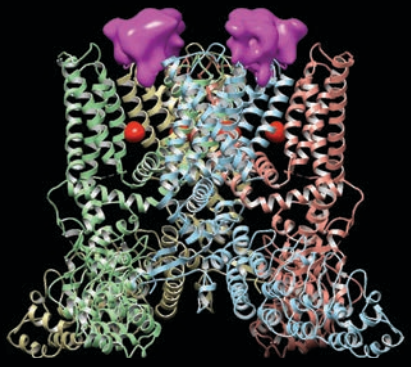


- DkTx binds to and trap TRPV1 in open state.
- DkTx targets the outer pore domain of TRPV1.

Julius Lab
Bohlen, et al. (2010) Cell

Structure of TRPV1-DkTx-RTX complex

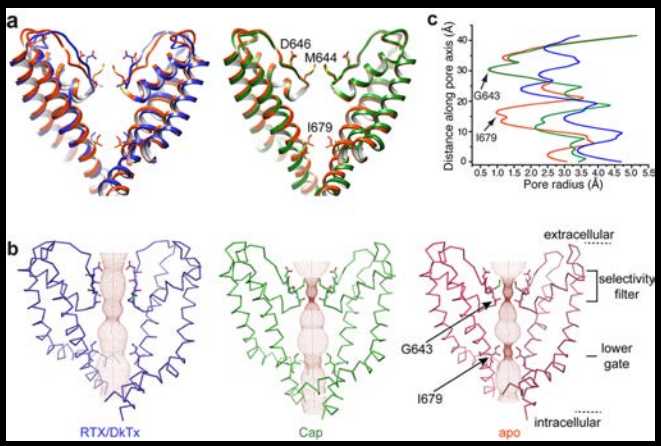
Binding sites of DkTx and RTX in TRPV1 channel.



Available and well-characterized pharmacological reagents made it possible to determine and relate the structures to its functional states.

Cao, Liao, et al
(2013) Nature

TRPV1 in three distinct pore conformations



* Binding of capsaicin opens lower gate;
* Binding of both DkTx and RTX opens both lower gate and upper selectivity filter;

Cao, Liao, et al (2013) Nature

How to handle difficult membrane proteins?

A typical flow chart of determining membrane protein structure by single particle cryo-EM

I. *Protein productions*: expression, purification, optimizations for biochemical conditions;

↓↑

II. *Single particle cryo-EM*: data acquisition, image processing -> final 3D reconstruction;

↓

III. *Model building*:
de novo model building and refinement;

↓

IV. *Manuscript writing*:
to publish or not yet to publish?
to publish a structure or to tell a story?

I. Protein productions

Recombinant proteins:

* Optimize protein expressions and purifications - test different orthologs: using Fluorescence-detection size exclusion chromatography (FSEC);

Endogenous proteins:

* Optimize purification protocol - to generate homogeneous and stable proteins;

-- It is critical to establish function assays to ensure that purified proteins are functional, and to provide means to for validating hypothesis generated from structures;

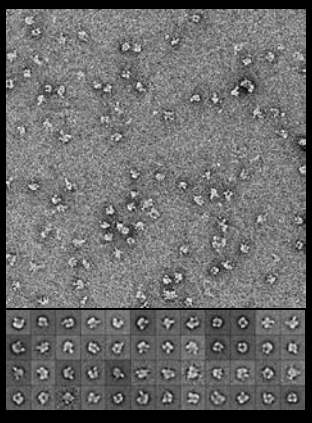
Optimizations:

* Optimize protein solubility, homogeneity and stability - test different detergents, protocols, etc, to generate homogeneous and stable proteins: combining SEC and negative stain EM;

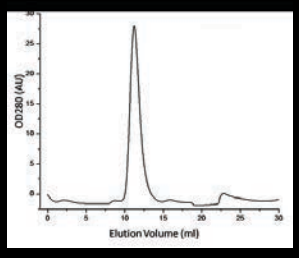
-- Our experience: Good SEC profile is required but not sufficient. In addition, every prep needs to be checked by negative stain EM;

Detergent solubilized rat TRPV1

* Rat TRPV1 was expressed in mammalian expression system (HEK293S);
* Solubilized and purified in the presence of detergent;



Negative stain EM



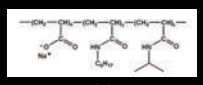
Size exclusion chromatography

* Tetramer MW: 73 kD x 4 = 292 kD;
* C4 symmetry

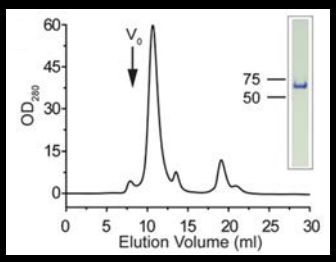
Liao, Cao, et al (2013) Nature

TRPV1 in amphipols

* Detergent was substituted with amphipols - leads to stable proteins for structural analysis.



A8-35



Size exclusion chromatography

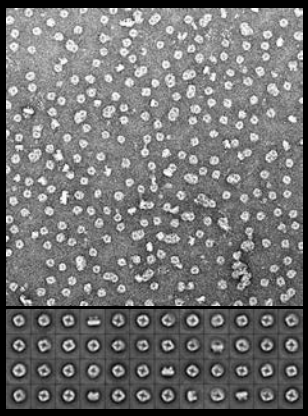
Tribet, et al. "Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions", PNAS (1996).

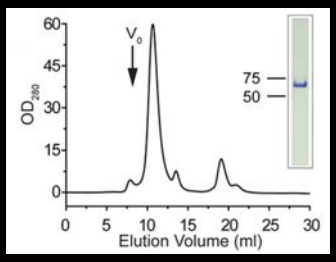
Althoff et al. "Arrangement of electron transport chain components in bovine mitochondrial supercomplex", EMBOJ (2011).

Liao, Cao, et al (2013) Nature

TRPV1 in amphipols

* Detergent was substituted with amphipols - leads to stable proteins for structural analysis.





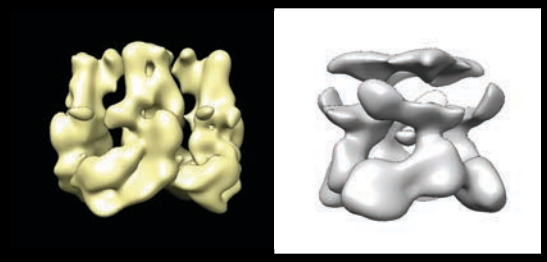
Size exclusion chromatography

Liao, Cao, et al (2013) Nature

Be Aware!

Not every sample goes to high resolution! Having biochemically well-characterized and well-behaved proteins is not a guaranty to produce high resolution structures.

* For membrane proteins: an often encountered problem is that the density of transmembrane domain is much weaker than soluble domain.



TRPV1 TRPV2

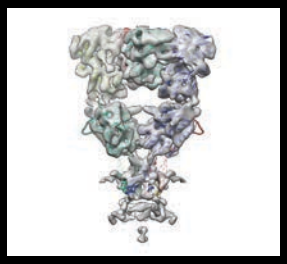
At low resolution

Be Aware!

Not every sample goes to high resolution! Having biochemically well-characterized and well-behaved proteins is not a guaranty to produce high resolution structures.

* For membrane proteins: an often encountered problem is that the density of transmembrane domain is much weaker than soluble domain.

γ-secretase, 4 ~ 5Å



GluR, 7.6Å

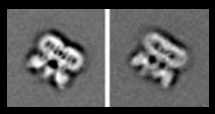
Or at high resolution

Be Aware!

Not every sample goes to high resolution! Having biochemically well-characterized and well-behaved proteins is not a guaranty to produce high resolution structures.

* For membrane proteins: an often encountered problem is that the density of transmembrane domain is much weaker than soluble domain.

The cause must be conformational heterogeneity in transmembrane domain, even when soluble domain is conformational homogeneous.



It is a good sign to see detailed features in transmembrane domain in 2D class averages.

Be Aware!

Not every sample goes to high resolution! Having biochemically well-characterized and well-behaved proteins is not a guaranty to produce high resolution structures.

* For membrane proteins: an often encountered problem is that the density of transmembrane domain is much weaker than soluble domain.

The cause must be conformational heterogeneity in transmembrane domain, even when soluble domain is conformational homogeneous.

- Potential solution A: go back to biochemistry!
  - Screen for different orthologous;
  - Using additives to stabilize transmembrane domains;
  - Lipid nanodisc to keep protein in lipid bi-layer environment;

- Potential solution B: sort out homogeneous conformation computationally!
  - Better classification algorithms to sort out different conformations.
  - (What is the physiologically relevant conformation?)

Importantly, better classification algorithms cannot substitute better biochemistry to produce better proteins to start with!

II. Cryo-EM data acquisition and processing

Good news: Data acquisition and processing is better established, more robust and more efficient;

* Direct detection camera is widely used, dose fractionation data acquisition is now “standard” procedure, motion corrections (both globe and local) are routinely applied;
* Automated or semi-automated data acquisitions enable novice users to collect high quality data;
* Automated particle picking making it possible to generate very large dataset within a reasonable time frame.
* Streamlined data processing enables novice user to determine high-resolution structures in a short period of times.

* Ideally, when all the resources are available, a typical cycle of determining a reconstruction, from data acquisition to final reconstruction, takes about a week.
* High efficiency made it possible to test samples prepared under different biochemical conditions.

Bad news: Very psychologically stressful!

III. Model building and refining

* *De novo* model building is new for many of us, and it is still very challenging for most structures.

- For maps with resolutions at borderline, or even some “near atomic resolution” maps: *de novo* model building is very challenging and time consuming.

We used to worry about “if my map is correct or not?”
- map validation methods: tilt pair, phase randomization, gold-standard refinement, etc.

Now we also worry (even more) about “if my atomic structure is correct, even my map is fine?”
- how to properly validate atomic structures?

Check and validation

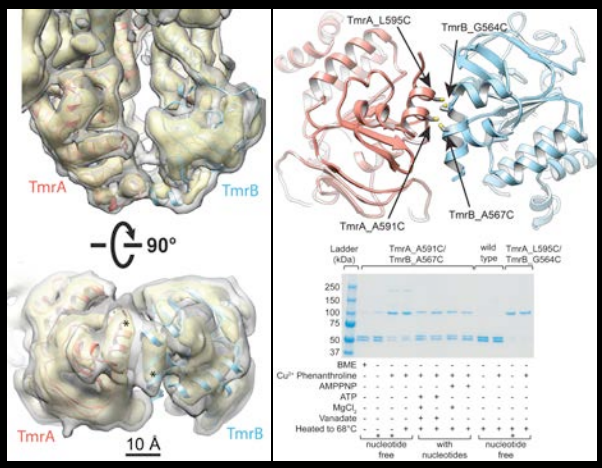
Computational refinement of atomic structures to ensure correct geometry

* How to adapt and modify existing X-ray crystallographic refinement methods for refining atomic structures built on cryo-EM maps?

* What is the definition of over-fitting and how to avoid it?

Validating atomic structures by cysteine cross-linkings.

Structure validation by cysteine cross-linking



* NBDs of TmrAB are in close contact at nucleotide free state;
* Cys cross-linking confirms such close contact in solution;

Kim, Wu, Tomasiak et al (2014) Nature

IV. Manuscript writing

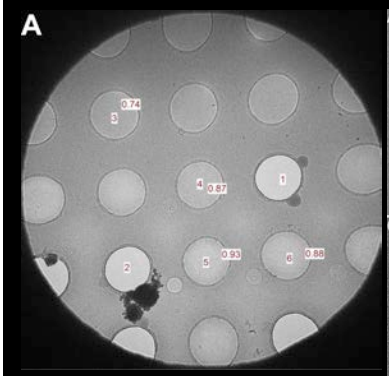
Question: what should we publish? A ***structure*** or a ***story***?

Lesson learnt from TRPV1: Available of well characterized pharmacological reagents enabled trapping protein in specific functional states.

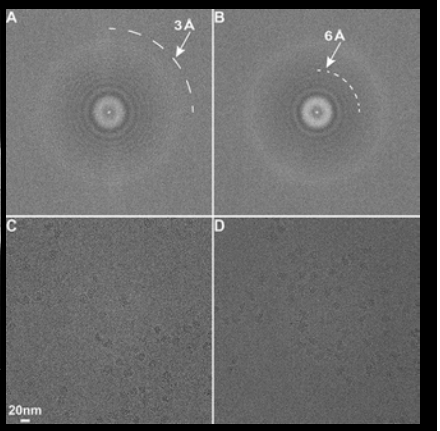
What are current limitations?

* Not every sample will go to high resolution. Why and how to change it?
  - Heterogeneity (both conformational and compositional)
  - Other limiting factors?

Influence of ice thickness



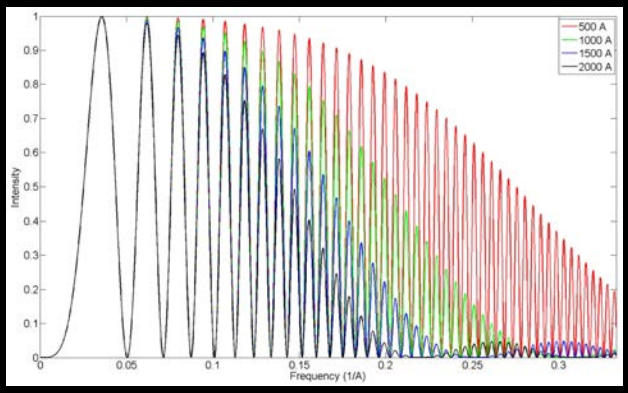
Estimated ice thickness: 3: ~903Å, 4: ~420Å, 5: ~220Å, and 6: ~380Å;



There is a strong correlation between visibility of Thon ring with the thickness of ice.

One possible explanation

Average of particles with different defocuses dampens the high-frequency Thon ring.



Simulation with 300kV, -2um defocus and different ice thickness.

Acknowledgement

Current members:
Shenping Wu, Jean-Paul Armache, David Bulkley, Yuan Gao, Ruchika, Shangyu Dang

Former members:
Yadong Yu
Homin Kim (KAIST, Korea)
Agustin Avila-Sakar (Gatan, Inc.)
Xueming Li (Tsinghua University, China)
Maofu Liao (Harvard Medical School)
David Booth, Kiyoshi Egami, Maggie Yang

K2 Project:
Agard Lab: David Agard, Shawn Zheng, Michael Braunfeld;

TRP ion channel:
Julius Lab: David Julius, Erhu Cao, Jeremiah Osteen, Candice Paulsen

ABC transporter:
Craik Lab: Charles Craik, JungMin Kim, Michael Winter
Stroud Lab: Robert Stroud, Thomas Tomasiak
Robert Tampe

Funding:
NSF: DBI-0960271 (MRI-consortium, K2 development);
NIH: R01GM098672 (EUREKA), P50GM082250 (HARC Center);
UCSF Program for Breakthrough Biomedical Research;