ISSN 1674-1056
http://cpb.iphy.ac.cn
http://iopscience.iop.org/cpb
中国物理 B
Chinese Physics
中国物理
Chinese Physics B

Structural biology revolution led by technical breakthroughs in cryo-electron microscopy

Chang-Cheng Yin(尹长城)
Citation: Chin. Phys. B . 2018, 27(5): 058703. **doi:** 10.1088/1674-1056/27/5/058703

Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Bio-macromolecular dynamic structures and functions, illustrated with DNA, antibody, and lipoprotein

Lu Gou(缑璐), Taoli Jin(金桃丽), Shuyu Chen(陈淑玉), Na Li(李娜), Dongxiao Hao(郝东晓), Shengli Zhang(张胜利), Lei Zhang(张磊)
Chin. Phys. B . 2018, 27(2): 028708. **doi:** 10.1088/1674-1056/27/2/028708

Recent technical advancements enabled atomic resolution CryoEM

Xueming Li(李雪明)
Chin. Phys. B . 2016, 25(1): 018710. **doi:** 10.1088/1674-1056/25/1/018710

Two-dimensional model of elastically coupled molecular motors

Zhang Hong-Wei,Wen Shu-Tang,Chen Gai-Rong,Li Yu-Xiao,Cao Zhong-Xing,Li Wei
Chin. Phys. B . 2012, 21(3): 038701. **doi:** 10.1088/1674-1056/21/3/038701

TOPICAL REVIEW — Electron microscopy methods for emergent materials and life sciences

Structural biology revolution led by technical breakthroughs in cryo-electron microscopy*

Chang-Cheng Yin(尹长城)†

Department of Biophysics, College of Basic Medicine, Peking University Health Science Center, Peking University, Beijing 100191, *China*

(Received 14 December 2017; revised manuscript received 22 February 2018; published online 26 April 2018)

Recent technical breakthroughs in cryo-electron microscopy (cryo-EM) revolutionized structural biology, which led to the 2017 Nobel Prize in chemistry being awarded to three scientists, Jacques Dubochet, Joachim Frank, and Richard Henderson, who made groundbreaking contributions to the development of cryo-EM. In this review, I will give a comprehensive review of the developmental history of cryo-EM, the technical aspects of the breakthrough in cryo-EM leading to the structural biology revolution, including electron microscopy, image recording devices and image processing algorithms, and the major scientific achievements by Chinese researchers employing cryo-EM, covering protein complexes involved in or related to gene expression and regulation, protein synthesis and degradation, membrane proteins, immunity, and viruses. Finally, I will give a perspective outlook on the development of cryo-EM in the future.

Keywords: cryo-electron microscopy (cryo-EM), structural biology, protein molecular machinery

PACS: 87.80.–y, 87.64.Ee, 87.15.B– **DOI:** 10.1088/1674-1056/27/5/058703

1. Introduction

Following the creative, cooperative efforts by hard-working scientists from different disciplines, including physics, computing science and structural biology, cryo-electron microscopy (cryo-EM) made a technical breakthrough and revolutionized structural biology, which led to the structural determination of a variety of macromolecular complexes ("protein machineries") and provided insights into their functions and working mechanisms. The 2017 Nobel Prize award to Jacques Dubochet, Joachim Frank and Richard Henderson marked a new era of structural biology. With regard to the original creative contributions by these three Nobel laureates to the development of cryo-EM and the reasons for the Nobel Prize award, readers should refer to a recent review.[1]

The breakthroughs in cryo-EM not only attracted great attention from structural biologists, but stimulated great interest among researchers in other fields, including biochemistry and molecular biology, cell biology, neurobiology, etc. Researchers would like to understand what cryo-EM is and what role it plays in structural biology. I have been studying and working in the MRC Laboratory of Molecular Biology (MRC-LMB) in Cambridge, the founding place of three-dimensional electron microscopy (3D EM) and have personal experience and understanding of the development history of cryo-EM. Here, I will give a comprehensive review of the development history of cryo-EM, the technical breakthroughs in cryo-EM leading to structural biology revolution and the major scientific achievements by Chinese researchers employing cryo-EM. The purpose of this review is to stimulate the interest of Chinese researchers in biological sciences to use cryo-EM as a powerful tool for solving structures of macromolecular complexes with great biological significance, and providing structural insights into their functions and working mechanisms.

2. Development history of cryo-EM

To understand cryo-EM, we first need to understand structural biology. Structural biology is a discipline that studies three-dimensional (3D) structures and structural transitions of macromolecules and their complexes at an atomic level by physical methods, hence providing structural insights into the functions and working mechanisms of macromolecules and their complexes. Structural biology originated from MRC-LMB (then in the Cavendish Laboratory, Cambridge University) in the 1940's, while Max Perutz and John Kendrew studied the structure of protein (hemoglobin), and James Watson and Francis Crick studied the structure of DNA by x-ray crystallographic methods. The successful structural determination of hemoglobin and the DNA double helix revolutionized biology—biology transformed from macroscopic and qualitative descriptions to microscopic and molecular mechanistic descriptions of biological phenomena, i.e., the molecular biology era. Max Perutz and John Kendrew were therefore awarded Nobel Prize in chemistry, and James Watson and Francis Crick were awarded the Nobel Prize in physiology/medicine in 1962.

Cryo-EM is a structural biology technique, in which an electron microscope is used to image macromolecules or their

*Project supported by the National Key Research and Development Program of China (Grant No. 2017YFA0504700) and the National Natural Science Foundation of China (Grant Nos. 31570732 and 31770785).

†Corresponding author. E-mail: ccyin@hsc.pku.edu.cn

© 2018 Chinese Physical Society and IOP Publishing Ltd

complexes frozen in a thin layer of vitreous ice, the images are processed and the 3D structures of macromolecules or their complexes are reconstructed from these images by computers (Fig. 1). The principle of 3D structure reconstruction from EM images was proposed by Aaron Klug and his colleagues at MRC-LMB in the 1960's. The principle is based on the Fourier transform and the central section theorem of Fourier transform. Central section theorem can be stated as the Fourier transform of a projection function of a 3D function equal to a central section of the Fourier transform of the 3D function. Another important nature of Fourier transform is that the inverse Fourier transform of a Fourier transform restores the original function. An EM image of an object is a projection of that object. Based on Fourier transform and central section theorem, Klug and his colleagues proposed the 3D structure reconstruction scheme shown in Fig. 2: first, taking images of macromolecules in different orientations; second, Fourier transforming the images, which are the central sections of the Fourier transform of the macromolecule structure; third, combining all the central sections of the Fourier transform and forming a gather; finally, inverse Fourier transforming the gather to restore the macromolecule structure (Fig. 2).[2] This scheme based on central section theorem laid the foundation for 3D structure reconstruction from EM images and subsequently developed structure determination approaches using electron microscopy, including electron crystallography and single particle analysis, which are essentially based on the scheme shown in Fig. 2. Due to this groundbreaking contribution in 3D EM, Klug was awarded the Nobel Prize in chemistry in 1982.

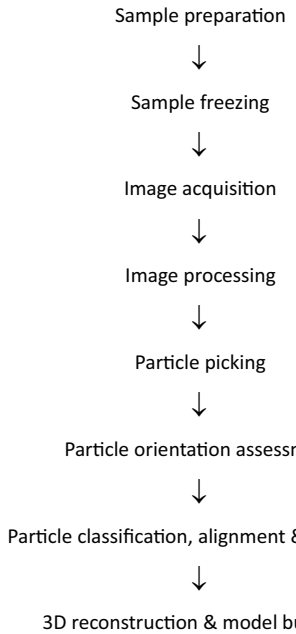


Fig. 1. The structure determination workflow of macromolecule by cryo-EM.

Although the 3D structure reconstruction scheme was proposed by Klug and his colleagues early in the 1960's, both the hardware for recording EM images and the software for image processing/3D reconstruction were not mature enough to determine the structures of macromolecules or their complexes to atomic resolution. Before 2013, only highly ordered arrays of protein samples (two-dimensional (2D) crystals) and icosahedral viruses could be determined to atomic resolution by cryo-EM. Since 2013, due to technical breakthroughs in both hardware and software, protein samples dispersed in solution (single particles) can now be determined to atomic resolution. The technical breakthroughs in cryo-EM brought about a structural biology revolution and hence led to a great leap forward in biological research.

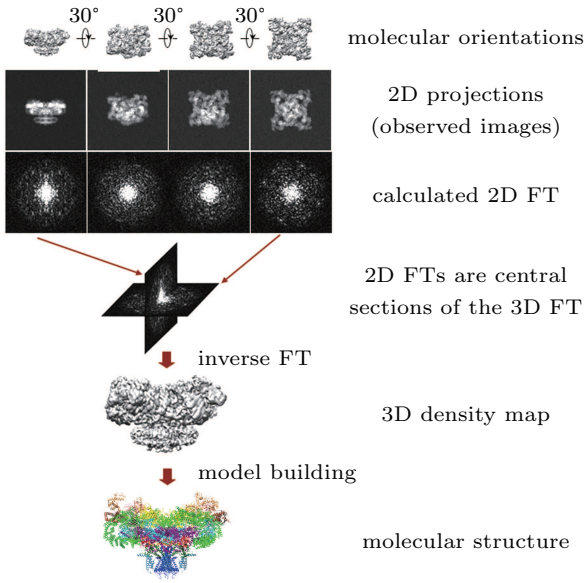


Fig. 2. (color online) Principle of 3D reconstruction of a macromolecule from EM images.[2]

3. Technical breakthroughs in cryo-EM

The technical breakthroughs in cryo-EM originated from breakthroughs in both hardware and software. In hardware, the electron microscope has led to great advances and image recording devices have made fundamental breakthroughs; in software, image processing/3D reconstruction algorithms have made great breakthrough. All breakthroughs will be described in detail in the following sections.

3.1. Electron microscope

In 2006, the FEI Company (now merged with the Thermo-Fisher Company) launched the state-of-the-art cryo-electron microscope—Titan Krios (Fig. 3). Compared with previous version, Titan Krios has technical advances in three aspects: first, Titan Krios has a three-level condenser system, which greatly improves the parallelism and coherence of emitting electrons and thus improves the resolution of EM images; second, Titan Krios is equipped with a grid-loading robot system, which can automatically select and load grids one-by-one from a liquid nitrogen-cooled cassette preloaded with 12 sample grids, thus greatly improving the sam-

ple grid loading efficiency and the success rate of finding a suitable and best sample grid for data collection; third, the loaded sample grid is isolated from the outside and cooled by a liquid nitrogen cooling system throughout a data collection session, which not only avoids contamination of the sample grid by ice, but also stabilizes the sample grid for a long time, thus greatly extending the data collection time (http://cbi.ibp.ac.cn/cbiweb/2012/1123/218.html). When both the electron microscope and the sample are in their best states, one might solve the structure of a target protein or its complex at atomic resolution in a single data collection session with just one sample grid.

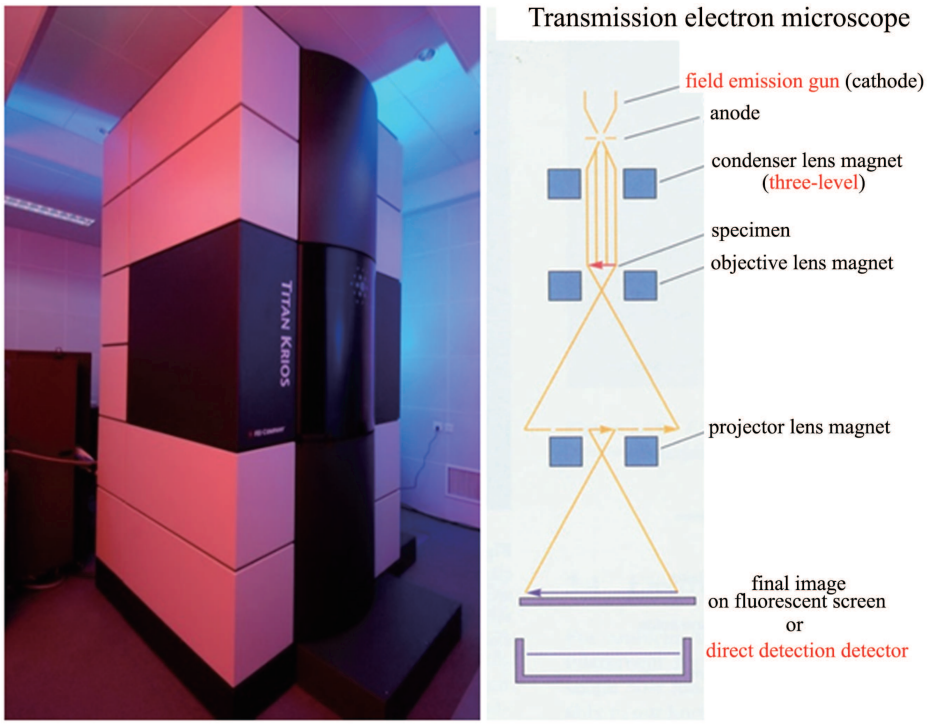


Fig. 3. (color online) Thermo-Fisher Titan Krios electron microscope (left); electronic-optical system of Titan Krios (right). The Titan Krios electron microscope is a field emission gun-type microscope, and has three-level condenser, while other electron microscopes have only two-level, and it is equipped with direct detection detector.

3.2. Image recording device—direct electron detector

Before 2013, EM images were recorded on electron-sensitive films or specifically-designed CCD for EM. Although the resolution recorded on electron-sensitive films is sufficiently high (atomic resolution), the number of images recorded from a single data collection session is limited due to the fixed number of the film box of an electron microscope (56 for FEI electron microscopes). The maximum number of images that could be recorded in a day is around 100. Besides, the recorded films needed to be developed, washed and dried, and scanned by a high-resolution image scanner to transform them into digital form, and then the images were input into a computer for image processing. This tedious process made the efficiency of structural determination from films extremely low, which limited its application in structure determination by cryo-EM. Specifically-designed CCD for EM records EM images in digital form and there was no limitation of the the images that could be recorded in a session, thus the efficiency of CCD in structure determination is much higher than film. However, during image recording the electron signals are transformed into photon signals in CCD, so the signal decays during transformation, thus the signal intensity decreases after transformation; in addition, the resolution of the images becomes deteriorated due to the point spread effect (point becomes circle); furthermore, as the frozen sample is very sensitive to electron radiation, cryo-EM images must be recorded in low-dose mode to minimize radiation damage by electrons, so a long expose time (normally 1 second) is used in recording images, the images recorded are normally blurred due to image drifting. The images recorded on CCD are of low signal-to-noise ratio and low resolution, normally lower than films.

In 1995, Richard Henderson wrote a theoretical essay on single-particle cryo-EM and predicted that single-particle cryo-EM was capable of achieving atomic resolution and pointed out that two key problems need to be overcome to achieve atomic resolution: the first is that the signal-to-noise ratio of the EM images needs to be improved, the second that the problem of image drifting of the EM images during image recording needs to be overcome.[3] Henderson's theoretical essay indicated the correct direction for the development of cryo-EM. In 2013, following the theoretical essay, Yifan Cheng and David Agard of UCSF, USA, recorded cryo-EM images on a direct electron detector (DED)/direct detection detector (DDD). DDD has great advantages over CCD: first, the DDD records electron signal directly without the transfor-

mation into photon signal, thus the original signal is maintained so the signal-to-noise ratio is substantially improved (the detective quantum efficiency (DQE) of a DDD is much higher than that for CCD and even higher than for electron-sensitive film (Fig. 4)[4]); second, the electron signal is directly recorded so the point spread effect is greatly reduced, hence the resolution of the recorded images is greatly improved; third, both the signal and the signal-to-noise ratio are high, so short a exposure time can be used to record EM images, thus a field could be imaged as consecutive frames (for instance, the FEI Falcon 2 DDD can record 24 frames per second, and the Gatan K2 Summit DDD can record 48 frames per second), and the image drifting during image recording could be corrected by aligning the recorded frames and averaging them, hence the blurring of the EM image caused by image drifting could be eliminated and the resolution of the recorded EM images is significantly improved. Under the circumstance at which both the electron microscope and the sample are in best states, one might solve the structure of a target protein or its complex to atomic resolution from images recorded on the DDD in a single data collection session with just one sample grid. Using the images recorded on a DDD, Yifan Cheng and his colleagues solved the near-atomic resolution structure of the TRVP1 channel by single-particle cryo-EM for the first time (Fig. 5).[5]

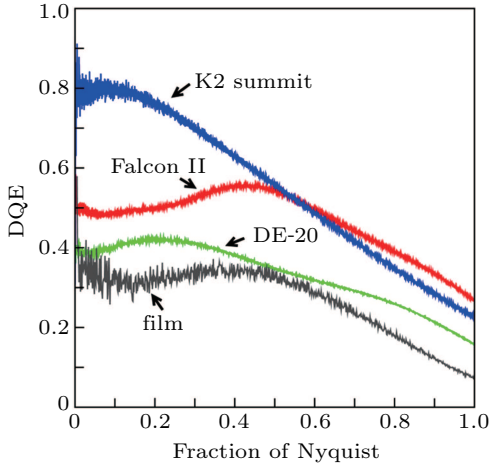


Fig. 4. (color online) The detective quantum efficiency (DQE) of direct detection detectors. Large DQE represents stronger detective electron signal for EM imaging.[4] Figure from Ref. [4] by courtesy of Richard Henderson (MRC-LMB, Cambridge, UK).

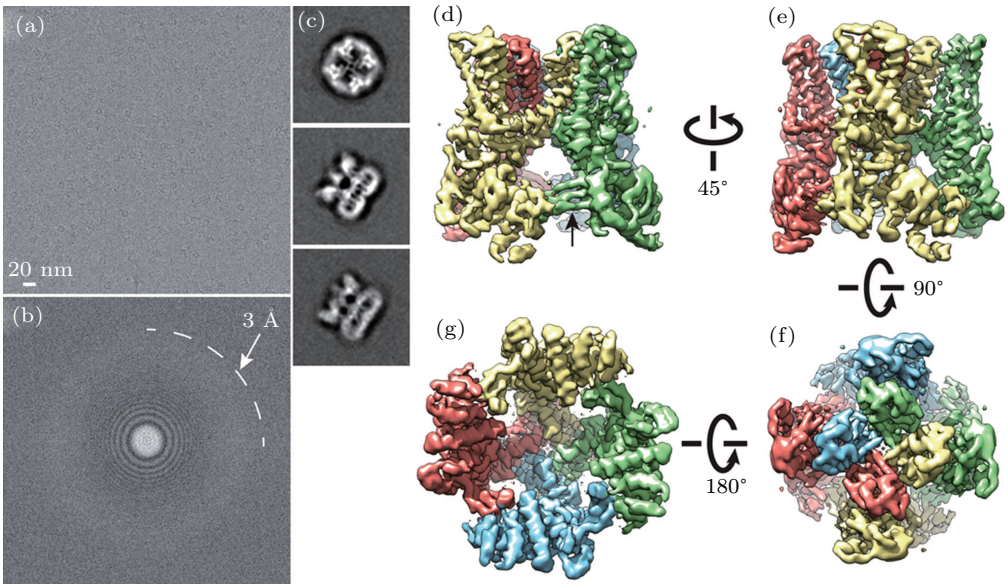


Fig. 5. (color online) The first near-atomic resolution structure (TRPV1) determined by single particle cryo-EM recorded by a DDD.[5] (a) An original photo; (b) the power spectrum (Fourier transform) of image after correction for image shift; (c) three typical projection views (2D averaged images) of TRPV1 with detailed structural information; (d)–(f) display the reconstructed density maps of TRPV1 in side (d and e), bottom (g), and top (f) views. The protein structural features (α-helices and β-sheets) are clearly identifiable. Figure from Ref. [5] by courtesy of Yifan Cheng (UCSF, USA).

3.3. Image processing and 3D reconstruction algorithms

Several image processing and 3D reconstruction software packages were commonly used in cryo-EM before 2013, including SPIDER, the first EM image processing and 3D reconstruction software developed by Joachim Frank and his colleagues,[6] IMAGIC, software developed by Marin van Heel,[7] FREALIGN, software developed by Nicola Grigorieff,[8] and EMAN, software developed by Steve Ludtke and his colleagues.[9] In cryo-EM, a purified macromolecule or its complex is used as the material. An assumption is that all the particle EM images are different projections of the same molecule in the same conformation. In solution, molecules would take different orientations due to Brownian motion; upon fast freezing, molecules would be frozen in random orientations. If the number of the same molecule in different orientations are sufficient (thousands to millions), then the gather of Fourier transforms of the particle images would cover the

3D Fourier space and thus give the 3D Fourier transform of the molecular structure, and the inverse Fourier transform of the 3D Fourier transform will give the reconstructed molecular structure. In image processing, the most critical step is the classification and orientation determination of each particle. In most image processing and 3D reconstruction software packages, particle classification is based on principal component analysis, and particle orientation is determined by cross-correlation of a particle image with different projections of a model molecule, with the orientation having the largest cross-correlation coefficient taken as particle orientation. As the signal-to-noise ratio of a cryo-EM image is very low (normally < 0.1), the cross-correlation coefficient of the particle image with the model is frequently less than 1, hence the orientation determined has error. Furthermore, macromolecules undergo dynamic conformational changes; multiple conformers might co-exist in solution. The assumption that all the particle EM images are different projections of the same molecule in the same conformation would mix up the different conformers, thus the reconstructed structure would be an averaged structure of different conformers. For the reasons mentioned above, the resolution of single-particle cryo-EM is limited to sub-nanometers without reaching atomic resolution.

In 2013, Sjors Scheres and his colleagues of MRC-LMB applied RELION, an image processing and 3D reconstruction software package,[10] to DDD-recorded cryo-EM images and obtained a near-atomic resolution structure of a ribosome complex. Meanwhile, they were able to sort out different conformers and complexes of the ribosome and solve their structures separately (Fig. 6).[11] In RELION, multiple conformers/complexes of a macromolecule are proposed to co-exist in solution, they are classified into different major classes according to maximum likelihood based on Bayes probability; particles in each major class are further classified into subclasses (corresponding to different projections of the same conformer/complex) again according to maximum-likelihood based on Bayes probability.[10] Such an image processing algorithm not only greatly improves the accuracy of classification and orientation assessment, so that subsequent 3D reconstruction to atomic resolution is achieved, but more importantly, such image processing algorithm separates different conformers/complexes and their structures can reconstructed simultaneously. Given that sufficient particle images (thousands to millions) are collected, different conformers/complexes of a target macromolecule might all be reconstructed to atomic resolution. If the time course of these different conformers/complexes could be determined, then the structural transitions of the target macromolecule during a functional process could be revealed, thus the working mechanism of the target macromolecule could be illuminated. The image processing and 3D reconstruction algorithms of RELION revolutionized structural biology, transforming it from a static to a dynamic and unifying structure and function of a target macromolecule. This landmark breakthrough in cryo-EM was therefore selected as one of the top ten breakthroughs in science in 2014 by Nature.[12]

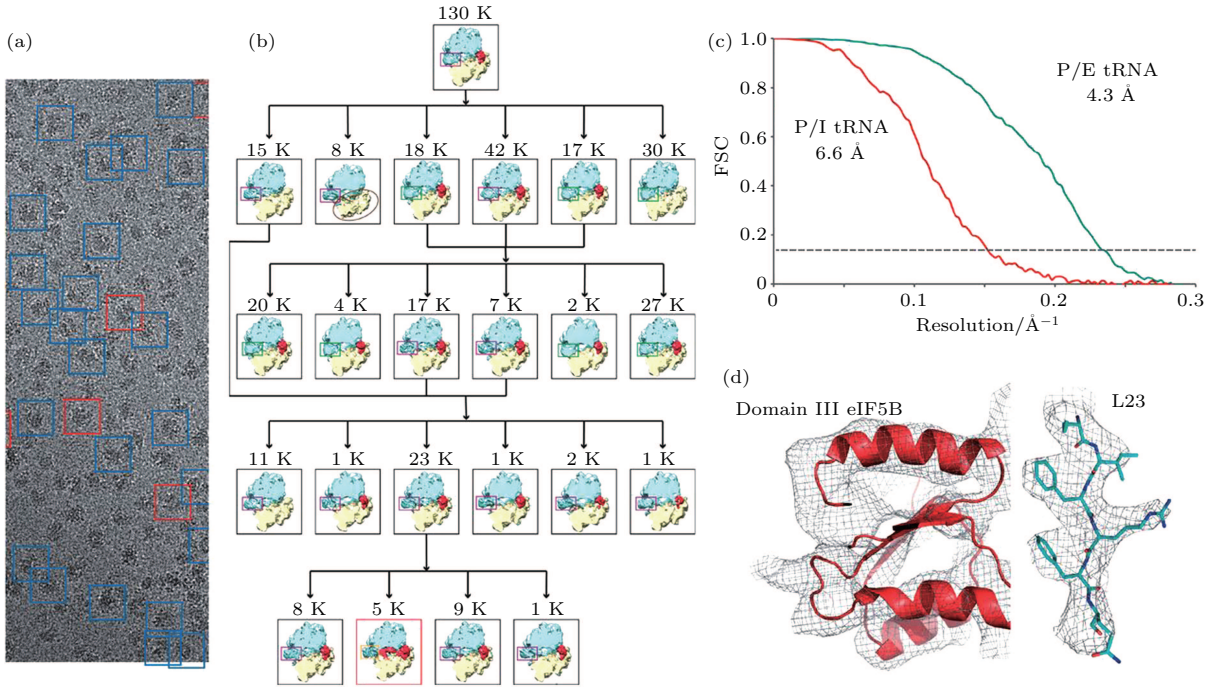


Fig. 6. (color online) The first near-atomic resolution structure (ribosome) determined by single particle cryo-EM by software RELION.[11] (a) is an original photo; (b) is the classification of different complexes/conformers existing in the original photos; (c) displays the resolution assessment by FSC 0.143 for two ribosome complexes (P/E tRNA and P/I tRNA complexes) with resolutions 4.3 Å and 6.6 Å, respectively; (d) display structural model (colored ribbons/sticks) fitting into EM density maps (grey mesh). Figure from Ref. [11] by courtesy of Sjors Scheres (MRC-LMB, Cambridge, UK)

4. Achievements of Chinese researchers employing cryo-EM

In 2005, the proposal by Zihe Rao and Fuchu He that China should strongly support protein science research to enhance international competitive capacity was adopted and the Chinese Government decided to set up the National Protein Research Facilities. In 2007, in the planning meeting on the National Protein Research Facilities (Beijing) organized by Fuchu He and Zhixin Wang, the proposal by Senfang Sui that the national protein research facilities should set up high-end cryo-EM facilities was accepted, and a consensus that the cryo-EM research platform should be taken as the key platform in the National Protein Research Facilities (Beijing) was reached. Under the leadership of Senfang Sui, I took part in the planning of cryo-EM facilities in the National Protein Research Facilities (Beijing) and proposed the configuration scheme of cryo-EM facilities. I also assisted the planning and the configuration scheme of cryo-EM facilities in the National Protein Research Facilities (Shanghai), Peking University, Zhejiang University, South University of Science and Technology of China, etc. The set-up of these high-end cryo-EM facilities in China laid the foundation for Chinese researchers to study important protein complexes employing cryo-EM, which had made plentiful and substantial achievements in recent years. Due to the limited space, here I only briefly review the major achievements by Chinese researchers employing cryo-EM together with biological/medical significances of these achievements and summarize the important achievements together with their biological/medical significances in Table 1. For those who are interested in a specific field, please refer to the original literatures cited in Table 1.

Table 1. Important protein complex structures determined by Chinese researchers employing cryo-EM.

Protein complex	Biological/medical function	Biological/medical significance	Research group	Reference
Gene expression/regulation complexes				
Spliceosome	splice introns out of pre-mRNA and joint the exons into mature mRNA	illuminate the working mechanism of spliceosome; reveal the complete working mechanism of central dogma	Yigong Shi	[13]–[20]
30-nm chromatin	higher order structure of chromatin; epigenetic regulation	reveal the folding mode of 30-nm chromatin; illuminate the mechanism of epigenetic regulation	Ping Zhu/Guohong Li	[21]
Acetylation enzyme of nucleosome NuA4	epigenetic regulation	reveal the mechanism of histone recognition	Ping Zhu/Zhucheng Chen	[22]
Chromatin remodeler Snf2	regulation of gene expression	reveal the mechanisms of DNA recognition and chromatin remodeling	Xueming Li/Zhucheng Chen	[23]
Transcription mediator	transcription regulation	reveal the assembly mode and working mechanism of mediator	Gang Cai	[24]
MCM helicase	replication regulation	reveal the assembly mode of MCM helicase	Ning Gao	[25]
Influenza RNA polymerase	Gene replication/transcription	illuminate the mechanism of replication of influenza gene	Yingfang Liu/Hongwei Wang	[26]
DNA recombinase RAD51	Gene homologous recombination	illuminate the mechanism of DNA homology search and strand-exchange	Hongwei Wang	[27]
Bacterial group II intron	Gene splicing and retrotranscription	illuminate the mechanisms of splicing and retrotranscription	Hongwei Wang	[28]
DNA-dependent protein kinase	DNA repair	illuminate the mechanisms of DNA recognition and protection	Yanhui Xu	[29]
RNA exosome complex	RNA processing	illuminate the mechanism of RNA-induced exosome activation	Hongwei Wang	[30]
Protein synthesis/degradation complexes				
Ribosome	protein synthesis	illuminate the mechanisms of protein synthesis and regulation	Ning Gao	[31]–[35]
Proteasome regulatory particle	protein degradation	illuminate the regulation mechanism of proteasome assembly	Youdong Mao	[36]
Protease HtrA	protein degradation	illuminate the mechanism of protein degradation	Senfang Sui	[37]
Membrane proteins				
Human γ-secretase	Alzheimer disease	reveal the enzyme action mechanism; guide drug design	Yigong Shi	[38]–[40]
Sodium channel Na(v)1.4-beta 1 complex	neural excitation	illuminate the mechanisms of channel gating and ion selectivity	Ning Yan	[41], [42]
Calcium channel Ca(v)1.1	calcium signaling; muscle cell excitation	illuminate the mechanisms of channel gating and ion selectivity	Ning Yan	[43], [44]

Table 1. (*Continued*).

Protein complex	Biological/medical function	Biological/medical significance	Research group	Reference
Calcium release channel RyR	calcium signaling; muscle cell excitation	illuminate the mechanisms of channel gating, regulation, and ion selectivity	Ning Yan Chang-Cheng Yin/Fei Sun	[45]–[47] [48]
Human lipid exporter ABCA1	lipid transport	illuminate the mechanisms of lipid binding and transport	Ning Yan	[49]
Ebola virus receptor	cholesterol transport; Ebola virus infection	illuminate the mechanisms of cholesterol trafficking and Ebola virus infection; guide drug design	Ning Yan/George F Gao	[50]
Mammalian respiratory supercomplex I1III2IV1	energy conversion	illuminate the mechanisms of electron transfer and energy conversion	Maojun Yang	[51]
Human respiratory megacomplex I2III2IV2	energy conversion	illuminate the mechanisms of electron transfer and energy conversion	Maojun Yang	[52]
Mammalian respirasome	energy conversion	reveal the assembly mode of mammalian respirasome	Maojun Yang	[53]
Mammalian mechano-sensitive channel Piezo1	mechanical force sensing	illuminate the mechanisms of mechanical force sensing and channel gating	Maojun Yang/Bailong Xiao/Ning Gao	[54]
Eukaryotic cyclic-nucleotide-gated channel	vision and olfaction	illuminate the mechanisms of channel gating and ion selectivity	Xueming Li/Jan Yang	[55]
Bacterial type II secretion channel GspD	substance transport	illuminate the mechanisms of channel gating and substance transport	Xueming Li/Yongqun Zhu	[56]
Photosystem II-LHCII supercomplex	photosynthesis	illuminate the mechanism of energy transfer	Zhenfeng Liu/Xinzheng Zhang	[57]
Plant C2S2M2-type PSII-LHCII supercomplex	photosynthesis	illuminate the mechanisms of light harvesting and energy transfer	Xinzheng Zhang/Zhenfeng Liu/Mei Li	[58]
Light-harvesting supercomplex	photosynthesis	illuminate the mechanisms of complex assembly and energy transfer	Senfang Sui	[59]
Membrane remodeling complex Vps4	membrane remodeling	illuminate the disassembly mechanism of endosomal sorting complex	Senfang Sui	[60]
SNAP-SNARE complex	membrane fusion	illuminate the disassembly mechanism of SNARE	Senfang Sui/Hongwei Wang	[61]
Immunity related complexes				
Inflammasome NLRC4	innate immunity	illuminate the assembly mechanism of NLRC4	Senfang Sui/Jijie Chai	[62]
Inflammasome complex NAIP-NLRC4	innate immunity	reveal the mechanism of signal amplification in inflammasomes	Youdong Mao	[63]
Viruses and related complexes				
Japanese encephalitis virus	cause encephalitis	identify structural elements that modulate stability and virulence; guide drug design	Zihe Rao/David Stuart	[64]
Hepatitis A virus-antibody complex	cause hepatitis	reveal the mechanism of neutralization; guide vaccine/drug design	Zihe Rao/David Stuart	[65]
Human Aichi virus	cause gastroenteritis in children	reveal receptor binding site; guide drug design	Zihe Rao/David Stuart	[66]
Ljungan virus	cause diabetes and myocarditis	illuminate the mechanism of virus genome encapsidation	Zihe Rao/David Stuart	[67]
Spike glycoproteins of coronaviruses	virus infection	illuminate the mechanism of receptor binding; guide vaccine/drug design	George F Gao/Xinzheng Zhang/Yi Shi	[68]
Spike glycoprotein of SARS-CoV virus	virus infection	illuminate the mechanism of receptor binding; guide vaccine/drug design	Ye Xiang/Xinquan Wang	[69]
Ebola virus glycoprotein-antibody complex	virus infection	reveal the mechanism of neutralization; guide vaccine design	Ye Xiang	[70]
dsRNA virus	architecture of dsRNA viruses	illuminate the mechanisms of the replication and transcription in dsRNA viruses	Hongrong Liu/Lingpeng Cheng	[71]

4.1. Complexes related to gene expression/regulation

4.1.1. Spliceosome

According to the "central dogma" of molecular biology proposed by Frances Crick, the genetic information stored in DNA must first be transcribed into mRNA, then the information encoded by mRNA is translated into protein, protein then performs the biological function that the gene encodes. Transcription transforms a DNA sequence into an RNA sequence. DNA contains both protein-encoding (called exon) and non-protein encoding (called intron) sequences. Transcription will

transform the full-length DNA sequence into the corresponding RNA sequence. Before translation, the introns must be removed and the remaining exons must be joined together to form a mature RNA called messenger RNA (mRNA), then the information can be translated into protein based on the mRNA template. The RNA processing is called RNA splicing and is performed by a giant protein enzyme complex called spliceosome. Due to a lack in detailed structural information, how the spliceosome complex is assembled, how it recognizes introns and removes them out of RNA, and how it joins the exons together to form mRNA remain unsolved mysteries. The working mechanisms of transcription and translation of the "central dogma" have been cracked. Roger Kornberg, who revealed the working mechanism of transcription, was awarded the 2006 Nobel Prize in chemistry and Venkatraman Ramakrishnan, Thomas Steitz and Ada Yonath, who illuminated the working mechanism of translation, were awarded 2009 Nobel Prize in chemistry. The working mechanism of splicing is the "missing link" of the "central dogma" between transcription and translation and a fundamental scientific problem remains to be cracked. From 2016, Yigong Shi and his colleagues in Tsinghua University solved a series of complex structures of the spliceosome existing in the splicing cycle by cryo-EM. From these structures, they deduced the working mechanisms of spliceosome complex assembly, intron recognition and removal, and exon joint together to form mRNA.

4.1.2. 30-nm chromatin

The collective length of genome of a human somatic cell would measure 2 meters though the nucleus where the genome is stored measures only several micrometres. How can a nucleus with a dimension of several micrometers accommodate DNA with a collective length of 2 meters? This is realized through four-step, multiple-level DNA folding into chromosome with the four-step folded structures corresponding to hierarchical organization of the chromatin: the primary structure is nucleosome, formed by DNA twining around histones; the secondary structure is 30 nm chromatin, formed by coiling of nucleosome; the tertiary structure is supersolenoid, formed by coiling of chromatin; the quaternary structure is the chromosome, formed by coiling of supersolenoid. Due a lack of detailed structural information, how DNA assembles into 30 nm chromatin and how this assembly is regulated remain unsolved mysteries and is a challenging problem in molecular biology. In 2014, Ping Zhu, Guohong Li and their colleagues at the Institute of Biophysics, Chinese Academy of Sciences solved the structure of reconstituted 30 nm chromatin by using cryo-EM. This structure not only revealed how the DNA assembles and folds in 30 nm chromatin, but more importantly, provided the structural basis for the epigenetic regulation of genes.

4.2. Complexes related to protein synthesis/degradation

Ribosome is the "factory" of cell for protein synthesis and thus is the source of biological activities of living things. The ribosome of a eukaryotic cell is composed of a 60S large subunit and a 30S small subunit. The complete ribosome translates the DNA information into protein based on its mRNA template. The assembly and function of the ribosome is precisely controlled by auxiliary proteins. Due to lacking detailed structural information, how auxiliary proteins control the assembly and function of the ribosome remains an unsolved mystery. In 2016, Ning Gao and his colleagues in Tsinghua University solved a series of complex structures of the yeast 60S pre-ribosome with an array of auxiliary proteins. They mapped 20 auxiliary proteins on the ribosome and determined their atomic resolution structures. These detailed structural information illuminated the mechanisms of protein synthesis and regulation by these auxiliary proteins.

4.3. Membrane proteins

Membrane proteins are the most important proteins in living things. A variety of biological activities, including neural excitation, substance transport, signal transduction, photosynthesis and energy conversion, are performed by membrane proteins. The structure of a membrane protein is the key to illuminating its working mechanism. Compared with water soluble proteins, membrane proteins are hard to crystallize and hence difficult to solve their structures by classical x-ray crystallographic methods. The breakthrough in single-particle cryo-EM provided a powerful tool for membrane protein structure determination. Since 2015, Chinese researchers solved a series of membrane proteins and/or their complexes involved in neural excitation, substance transport, signal transduction, photosynthesis and energy conversion by cryo-EM, which illuminated their working mechanisms. The major achievements include: human γ-secretase, solved by Yigong Shi (Tsinghua University) and his colleagues; sodium channel Na(v)1.4-beta 1 complex, calcium channel Ca(v)1.1, calcium release channels RyR1 and RyR2, human lipid transporter ABCA1, human lipid exporter related to Ebola virus infection, all solved by Ning Yan (Tsinghua University) and her colleagues; mammalian and human respiratory super complexes I1III2IV1 and I2III2IV2, mammalian respirasome, mammalian mechanosensitive channel Piezo1, all solved by Maojun Yang (Tsinghua University) and his colleagues; photosystem II-LHCII supercomplex and plant C2S2M2-type PSII-LHCII supercomplex, solved by Xinzheng Zhang/Zhenfeng Lu (Institute of Biophysics, Chinese Academy of Sciences) and their colleagues; eukaryotic cyclic-nucleotide-gated channel and bacterial type II secretion channel GspD, solved by Xueming Li (Tsinghua University) and his colleagues; algae light-harvesting supercomplex phycobilisome, solved by Senfang

Sui (Tsinghua University) and his colleagues.

4.4. Viruses and related complexes

Viruses are the pathogens leading to serious diseases, including influenza, hepatitis, encephalitis, AIDS, etc. The structures of viruses are not only the key to the mechanistic understanding of how viruses infect their host cells and how viruses proliferate themselves in host cells, but may also provide the foundation for structure-based rational design of anti-virus drugs or vaccines. Because viruses are much bigger than proteins and easily recognized in cryo-EM images, they are ideal objects for single-particle cryo-EM. Since 2015, Chinese researchers solved the structures of a series of viruses and/or their complexes related to serious diseases. These structures illuminate working mechanisms of virus infection to and proliferation in host cells and provide guide to the rational design of anti-virus drugs and vaccines. The important achievements include: Japanese encephalitis virus, hepatitis A virus-antibody complex, human Aichi virus, Ljungan virus, all solved by Zihe Rao (Tsinghua University/Institute of Biophysics, Chinese Academy of Sciences) and his colleagues; spike glycoproteins of coronaviruses, solved by Xinzheng Zhang/Fu Gao (Institute of Biophysics/Institute of Microbiology, Chinese Academy of Sciences) and their colleagues; spike glycoprotein of SARS-CoV virus and Ebola virus glycoprotein-antibody complex, solved by Ye Xiang (Tsinghua University) and his colleagues.

5. Perspective outlook

The technical breakthroughs in cryo-EM have led to a revolution in structural biology. This has resulted in the breakthroughs in the structure determination of protein "molecular machineries" and the cracking of their working mechanisms. More and more universities and research institutions in China have realized the potential of cryo-EM in biological research and are setting up high-end cryo-EM facilities. It can be expected that more important achievements will be made in the coming years by Chinese researchers employing cryo-EM. Currently, the achievements are confined *in vitro* to purified protein molecular machineries. The new trend in cryo-EM development would be "*in situ* structural biology" to study structures and interactions of molecular machineries *in situ* in cells and tissues and correlate structures and interactions with biological functions, thus unifying structure and function in the cellular context. Chinese researchers have fully realized the significance of technical breakthroughs in cryo-EM to the structural biology revolution, and have devoted themselves to the development of *in situ* cryo-EM.[72–83] We hope Chinese researchers will make breakthroughs in the development of *in situ* cryo-EM and make a groundbreaking contribution to the revolution of *in situ* structural biology.

Acknowledgment

I would like to thank Drs. Richard Henderson and Sjors Scheres (MRC Laboratory of Molecular Biology, Cambridge, UK) and Yifan Cheng (UCSF, USA) for providing original figures used in this review.

References

[1] Yin C C 2017 *Chin. J. Biochem. Mol. Biol.* **33** 979
[2] De Rosier D J and Klug A 1968 *Nature* **217** 130
[3] Henderson R 1995 *Q. Rev. Biophys.* **28** 171
[4] McMullan G, Faruqi A R, Clare D and Henderson R 2014 *Ultramicroscopy* **147** 156
[5] Liao M, Cao E, Julius D and Cheng Y 2013 *Nature* **504** 107
[6] Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M and Leith A 1996 *J. Struct. Biol.* **116** 190
[7] van Heel M, Harauz G, Orlova EV, Schmidt R and Schatz M 1996 *J. Struct. Biol.* **116** 17
[8] Grigorieff N 2007 *J. Struct. Biol.* **157** 117
[9] Ludtke S J, Baldwin P R and Chiu W 1999 *J. Struct. Biol.* **128** 82
[10] Scheres S H 2012 *J. Struct. Biol.* **180** 519
[11] Fernández I S, Bai X C, Hussain T, Kelley A C, Lorsch J R, Ramakrishnan V and Scheres S H W 2013 *Science* **342** 1240585
[12] Callaway E 2014 *Nature* **516** 318
[13] Yan C, Hang J, Wan R, Huang M, Wong C C and Shi Y 2015 *Science* **349** 1182
[14] Wan R, Yan C, Bai R, Wang L, Huang M, Wong C C and Shi Y 2016 *Science* **351** 466
[15] Wan R, Yan C, Bai R, Huang G and Shi Y 2016 *Science* **353** 895
[16] Yan C, Wan R, Bai R, Huang G and Shi Y 2016 *Science* **353** 904
[17] Yan C, Wan R, Bai R, Huang G and Shi Y 2017 *Science* **355** 149
[18] Zhang X, Yan C, Hang J, Finci LI, Lei J and Shi Y 2017 *Cell* **169** 918
[19] Wan R, Yan C, Bai R, Lei J and Shi Y 2017 *Cell* **171** 120
[20] Bai R, Yan C, Wan R, Lei J and Shi Y 2017 *Cell* **171** 1589
[21] Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu RM, Zhu P and Li G 2014 *Science* **344** 376
[22] Xu P, Li C, Chen Z, Jiang S, Fan S, Wang J, Dai J, Zhu P and Chen Z 2016 *Mol. Cell* **63** 965
[23] Liu X, Li M, Xia X, Li X and Chen Z 2017 *Nature* **544** 440
[24] Wang X, Sun Q, Ding Z, Ji J, Wang J, Kong X, Yang J and Cai G 2014 *Cell Res.* **24** 796
[25] Li N, Zhai Y, Zhang Y, Li W, Yang M, Lei J, Tye B K and Gao N 2015 *Nature* **524** 186
[26] Chang S, Sun D, Liang H, *et al.* 2015 *Mol. Cell* **57** 925
[27] Xu J, Zhao L, Xu Y, Zhao W, Sung P and Wang H W 2017 *Nat. Struct. Mol. Biol.* **24** 40
[28] Qu G, Kaushal P S, Wang J, Shigematsu H, Piazza C L, Agrawal R K, Belfort M and Wang H W 2016 *Nat. Struct. Mol. Biol.* **23** 549
[29] Yin X, Liu M, Tian Y, Wang J and Xu Y 2017 *Cell Res.* **27** 1341
[30] Liu J J, Niu C Y, Wu Y, Tan D, Wang Y, Ye M D, Liu Y, Zhao W, Zhou K, Liu Q S, Dai J, Yang X, Dong M Q, Huang N and Wang H W 2016 *Cell Res.* **26** 822
[31] Feng B, Mandava C S, Guo Q, Wang J, Cao W, Li N, Zhang Y, Zhang Y, Wang Z, Wu J, Sanyal S, Lei J and Gao N 2014 *PLoS Biol.* **12** e1001866
[32] Wu S, Tutuncuoglu B, Yan K, Brown H, Zhang Y, Tan D, Gamalinda M, Yuan Y, Li Z, Jakovljevic J, Ma C, Lei J, Dong M Q, Woolford JL Jr and Gao N 2016 *Nature* **534** 133
[33] Ma C, Kurita D, Li N, Chen Y, Himeno H and Gao N 2017 *Nature* **541** 550
[34] Ma C, Wu S, Li N, Chen Y, Yan K, Li Z, Zheng L, Lei J, Woolford J L Jr and Gao N 2017 *Nat. Struct. Mol. Biol.* **24** 214
[35] Li Z, Guo Q, Zheng L, Ji Y, Xie Y T, Lai D H, Lun Z R, Suo X and Gao N 2017 *Cell Res.* **27** 1275
[36] Lu Y, Wu J, Dong Y, Chen S, Sun S, Ma Y B, Ouyang Q, Finley D, Kirschner M W and Mao Y 2017 *Mol. Cell* **67** 322
[37] Zhang L, Wang X, Fan F, Wang H W, Wang J, Li X and Sui S F 2017 *Cell Res.* **27** 582
[38] Lu P, Bai X C, Ma D, Xie T, Yan C, Sun L, Yang G, Zhao Y, Zhou R, Scheres S H W and Shi Y 2014 *Nature* **512** 166

[39] Sun L, Zhao L, Yang G, Yan C, Zhou R, Zhou X, Xie T, Zhao Y, Wu S, Li X and Shi Y 2015 *Proc. Natl. Acad. Sci. U S A* **112** 6003

[40] Bai X C, Yan C, Yang G, Lu P, Ma D, Sun L, Zhou R, Scheres S H W and Shi Y 2015 *Nature* **525** 212

[41] Shen H, Zhou Q, Pan X, Li Z, Wu J and Yan N 2017 *Science* **355** pii: eaal4326

[42] Yan Z, Zhou Q, Wang L, Wu J, Zhao Y, Huang G, Peng W, Shen H, Lei J and Yan N 2017 *Cell* **170** 470

[43] Wu J, Yan Z, Li Z, Yan C, Lu S, Dong M and Yan N 2015 *Science* **350** aad2395.

[44] Wu J, Yan Z, Li Z, Qian X, Lu S, Dong M, Zhou Q and Yan N 2016 *Nature* **537** 191

[45] Yan Z, Bai X, Yan C, Wu J, Li Z, Xie T, Peng W, Yin C, Li X, Scheres S H W, Shi Y and Yan N 2015 *Nature* **517** 50

[46] Peng W, Shen H, Wu J, Guo W, Pan X, Wang R, Chen S R and Yan N 2016 *Science* **354** pii: aah5324

[47] Bai XC, Yan Z, Wu J, Li Z and Yan N 2016 *Cell Res.* **26** 995

[48] Wei R, Wang X, Zhang Y, Mukherjee S, Zhang L, Chen Q, Huang X, Jing S, Liu C, Li S, Wang G, Xu Y, Zhu S, Williams A J, Sun F and Yin C C 2016 *Cell Res.* **26** 977

[49] Qian H, Zhao X, Cao P, Lei J, Yan N and Gong X 2017 *Cell* **169** 1228

[50] Gong X, Qian H, Zhou X, Wu J, Wan T, Cao P, Huang W, Zhao X, Wang X, Wang P, Shi Y, Gao G F, Zhou Q and Yan N 2016 *Cell* **165** 1467

[51] Wu M, Gu J, Guo R, Huang Y and Yang M 2016 *Cell* **167** 1598

[52] Guo R, Zong S, Wu M, Gu J and Yang M 2017 *Cell* **170** 1247

[53] Gu J, Wu M, Guo R, Yan K, Lei J, Gao N and Yang M 2016 *Nature* **537** 639

[54] Ge J, Li W, Zhao Q, Li N, Chen M, Zhi P, Li R, Gao N, Xiao B and Yang M 2015 *Nature* **527** 64

[55] Li M, Zhou X, Wang S, Michailidis I, Gong Y, Su D, Li H, Li X and Yang J 2017 *Nature* **542** 60

[56] Yan Z, Yin M, Xu D, Zhu Y and Li X 2017 *Nat. Struct. Mol. Biol.* **24** 177

[57] Wei X, Su X, Cao P, Liu X, Chang W, Li M, Zhang X and Liu Z 2016 *Nature* **534** 69

[58] Su X, Ma J, Wei X, Cao P, Zhu D, Chang W, Liu Z, Zhang X and Li M 2017 *Science* **357** 815

[59] Zhang J, Ma J, Liu D, Qin S, Sun S, Zhao J and Sui S F 2017 *Nature* **551** 57

[60] Sun S, Li L, Yang F, Wang X, Fan F, Yang M, Chen C, Li X, Wang H W and Sui S F 2017 *Nat. Commun.* **8** 16064

[61] Zhou Q, Huang X, Sun S, Li X, Wang H W and Sui S F 2015 *Cell Res.* **25** 551

[62] Hu Z, Zhou Q, Zhang C, Fan S, Cheng W, Zhao Y, Shao F, Wang H W, Sui S F and Chai J 2015 *Science* **350** 399

[63] Zhang L, Chen S, Ruan J, Wu J, Tong A B, Yin Q, Li Y, David L, Lu A, Wang W L, Marks C, Ouyang Q, Zhang X, Mao Y and Wu H 2015 *Science* **350** 404

[64] Wang X, Li SH, Zhu L, Nian Q G, Yuan S, Gao Q, Hu Z, Ye Q, Li X F, Xie D Y, Shaw N, Wang J, Walter T S, Huiskonen J T, Fry E E, Qin C F, Stuart D I and Rao Z 2017 *Nat. Commun.* **8** 14

[65] Wang X, Zhu L, Dang M, Hu Z, Gao Q, Yuan S, Sun Y, Zhang B, Ren J, Kotecha A, Walter T S, Wang J, Fry E E, Stuart D I and Rao Z 2017 *Proc. Natl. Acad. Sci. USA* **114** 770

[66] Zhu L, Wang X, Ren J, Kotecha A, Walter T S, Yuan S, Yamashita T, Tuthill T J, Fry E E, Rao Z and Stuart D I 2016 *Nat. Microbiol.* **1** 16150

[67] Zhu L, Wang X, Ren J, Porta C, Wenham H, Ekström J O, Panjwani A, Knowles N J, Kotecha A, Siebert C A, Lindberg A M, Fry E E, Rao Z, Tuthill T J and Stuart D I 2015 *Nat. Commun.* **6** 8316

[68] Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, Lu G, Wu Y, Yan J, Shi Y, Zhang X and Gao G F 2017 *Nat. Commun.* **8** 15092

[69] Gui M, Song W, Zhou H, Xu J, Chen S, Xiang Y and Wang X 2017 *Cell Res.* **27** 119

[70] Misasi J, Gilman M S, Kanekiyo M, Gui M, Cagigi A, Mulangu S, Corti D, Ledgerwood J E, Lanzavecchia A, Cunningham J, Muyembe-Tamfun J J, Baxa U, Graham B S, Xiang Y, Sullivan N J and McLellan J S 2016 *Science* **351** 1343

[71] Liu H and Cheng L 2015 *Science* **349** 1347

[72] Wei D Y and Yin C C 2010 *J. Struct. Biol.* **172** 211

[73] Han R, Zhang F, Wan X, Fernández J J, Sun F and Liu Z 2014 *J. Struct. Biol.* **186** 167

[74] Han R, Wang L, Liu Z, Sun F and Zhang F 2015 *J. Struct. Biol.* **192** 403

[75] Liu B, Xue Y, Zhao W, Chen Y, Fan C, Gu L, Zhang Y, Zhang X, Sun L, Huang X, Ding W, Sun F, Ji W and Xu T 2015 *Sci. Rep.* **5** 13017.

[76] Zhang J, Ji G, Huang X, Xu W and Sun F 2016 *J. Struct. Biol.* **194** 218

[77] Chen Y, Zhang Y, Zhang K, Deng Y, Wang S, Zhang F and Sun F 2016 *J. Struct. Biol.* **195** 49

[78] Deng Y, Chen Y, Zhang Y, Wang S, Zhang F and Sun F 2016 *J. Struct. Biol.* **195** 100

[79] Han R, Wan X, Wang Z, Hao Y, Zhang J, Chen Y, Gao X, Liu Z, Ren F, Sun F and Zhang F 2017 *J. Struct. Biol.* **199** 196

[80] Chen Y, Wang Z, Zhang J, Li L, Wan X, Sun F and Zhang F 2017 *Biophys. Rep.* **3** 36

[81] Fan X, Zhao L, Liu C, Zhang J C, Fan K, Yan X, Peng H L, Lei J and Wang H W 2017 *Structure* **25** 1623

[82] Li S, Ji G, Shi Y, Klausen L H, Niu T, Wang S, Huang X, Ding W, Zhang X, Dong M, Xu W and Sun F 2018 *J. Struct. Biol.* **201** 63

[83] Wang Z, Chen Y, Zhang J, Li L, Wan X, Liu Z, Sun F and Zhang F 2017 *J. Comput. Biol.*