The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

2169-1

Conference on Molecular Aspects of Cell Biology: A Perspective from Computational Physics

11 - 15 October 2010

Tackling Large Macromolecular Complexes: The Interplay of Electron Microscopy and Protein Crystallography

S. ONESTI
Sincrotrone Trieste S.C.p.A.
Trieste
Italy

Imperial College London

elettra

Tackling large macromolecular complexes: the interplay between electron microscopy and protein crystallography.

A case study: structural studies of MCM helicases

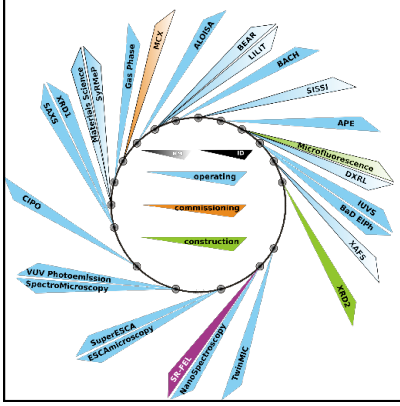
Silvia Onesti
silvia.onesti@elettra.trieste.it

ELETTRA

A storage ring operating between 2.0 and 2.4 GeV produces synchrotron light with wavelengths ranging from infrared to X-rays.

A new fourth generation light source based on Free Electron Laser is now being developed (FERMI@Elettra)

Beamlines



- 22 beamlines in operation and open to external users
- 1 beamline under commissioning
- 2 beamlines under construction
- 1 FEL beamline

Macromolecular crystallography beamlines:
- XRD1
- XRD2

Overview

- An introduction to single particle electron microscopy: strengths and weaknesses
- The biological problem:
  - DNA replication in eukaryotes
  - MCM helicases
  - Structural studies of MCM helicases
  - How does it work
  - How is it loaded onto DNA
- A comparison between macromolecular crystallography (MX) and single particle electron microscopy (EM)

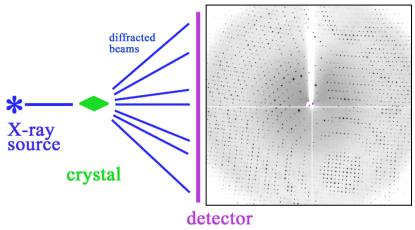
Overview

- An introduction to single particle electron microscopy: strengths and weaknesses
- The biological problem:
  - DNA replication in eukaryotes
  - MCM helicases
  - Structural studies of MCM helicases
  - How does it work
  - How is it loaded onto DNA
- A comparison between macromolecular crystallography (MX) and single particle electron microscopy (EM)

Structural biology

Macromolecular crystallography

If a X-ray beam is directed at a crystalline sample it will be scattered by the electrons in the sample and produce a diffraction pattern with discrete spots.



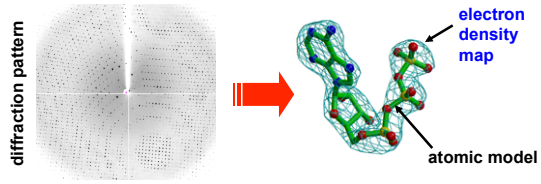
Why a crystal?
a single molecule would produce a diffraction pattern so weak as to be impossible to measure!

In the case of large molecules such as proteins, these patterns can be extremely complex, with thousands of spots.

Structural biology

Macromolecular crystallography

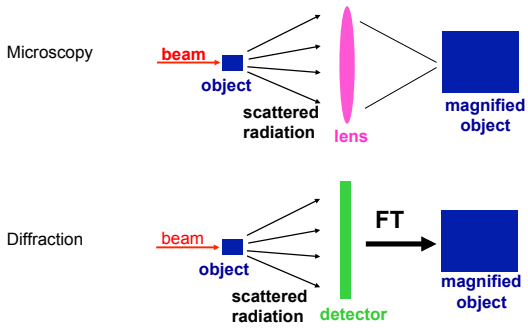
The X-rays are scattered by the electrons in the molecules. From the diffraction pattern we can work backwards and determine where the electrons responsible for the scattering are located. We obtain an **electron density map**.



If we know where the electrons are, we can deduce where the atoms are; the electron density map can be interpreted in terms of **atomic model**.

Structural biology

Microscopy vs diffraction



Single particle electron microscopy

Sample preparation

Samples for EM have to be carefully prepared:

- they have to be exposed to high vacuum and therefore fixed with special chemicals or frozen

- have to be prepared in extremely thin sections since electrons have limited penetrating power

- samples are often exposed to heavy metals since contrast depends on the atomic number (staining)

Here we focus on EM applied to single molecules

Single particle electron microscopy

Sample preparation: negative staining vs cryo

Negative staining
(usually 0.5% uranyl acetate)

- Simple procedure
- Quick to check samples
- High contrast
- Low radiation damage
- Can do tilt series
- Dehydration
- Artefacts due to uneven deposition of stain
- Distortions
- Preferential orientation?
- Low resolution (20 Å)

Cryo

- More complex preparation
- Time consuming
- Low contrast
- Radiation damage→low dose
- Cannot tilt
- Native, hydrated state
- 3D structure preserved
- In solution, no distortion
- Random orientation
- Rapid freezing can trap transient states
- High resolution (up to 3Å?)

Single particle electron microscopy

The problems

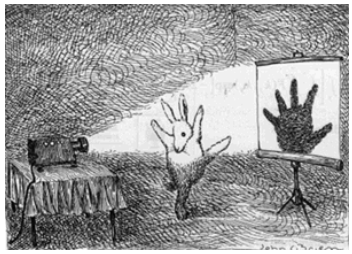
- image is 2D projection of original 3D object
- 3D structure can be determined from a set of views at different orientations
- radiation damage is the ultimate limit on resolution - to avoid destroying the sample, one uses very low doses, obtaining very noisy images

Therefore we have two problems

- getting a signal out of very noisy images
- getting a 3D structure out of 2D projections

Single particle electron microscopy

From 2D projections to 3D models



(The New Yorker, 1991)

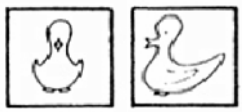
Obviously one projection is not sufficient to infer the structure of an object...

Single particle electron microscopy

From 2D projections to 3D models

From a number of projections in different directions we can reconstitute the object - the more projections we have, the more details we can get.

3D object



2D projections

The difficult step in 3D image processing is to determine the orientation angles (Euler angles) for each projection image.

Single particle electron microscopy

Dealing with noise - averaging

A single protein molecule gives only a weak and ill-defined image. Combine the information from many molecules so as to average out the random errors in the single images.

Averaging large number of particles massively increases the signal/ noise. However, we have to make sure we are averaging views with the same orientation!

3D reconstruction

Class averages are analysed to define their relative orientation (not easy!) and Euler angles are assigned to each class average. Once a preliminary solution is found, the views are combined to generate a 3D object. The 3D map is then re-projected in the directions indicated by the Euler angles, and each re-projection is compared to the corresponding class average.

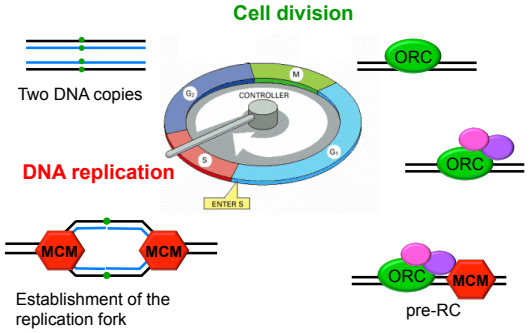
The process is cycled and the 3D model refined until convergence.

Overview

- An introduction to single particle electron microscopy: strengths and weaknesses
- The biological problem:
  - DNA replication in eukaryotes
  - MCM helicases
  - Structural studies of MCM helicases
  - How does it work
  - How is it loaded onto DNA
- A comparison between macromolecular crystallography (MX) and single particle electron microscopy (EM)

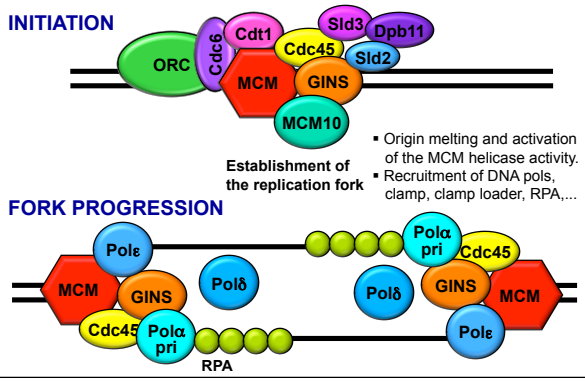
DNA replication

Cell cycle control of DNA replication



DNA replication

The initiation of DNA replication

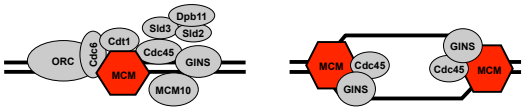


MCM helicases

Key players in DNA replication

A family of proteins conserved in all eukaryotic cells

Dual role in DNA replication:



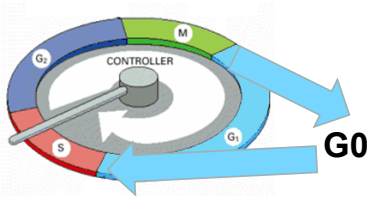
INITIATION **PROGRESSION**

MCM helicases

DNA replication & cancer

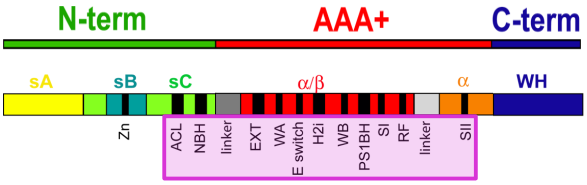
Most of the differentiated cells in our bodies are not in the cell cycle, but in a **quiescent state (G0)**. On the contrary proliferating cells are actively undergoing cell cycle. Hence DNA replication in adults is often connected with unwanted proliferation (a hallmark of cancer).

MCM helicases are only present in replicating cells, and can therefore be used as tumor markers and possible drug targets for cancer.



MCM helicases

MCM primary sequence



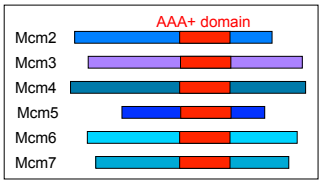
The AAA+ is the "engine" of the protein – uses the energy derived from hydrolysis of ATP to drive the unwinding of double-stranded DNA

Numerous elements have been biochemically characterised as being functionally important (trasducing chemical energy into mechanical work

MCM helicases

Eukaryotic & archeal MCM proteins

Eukaryotic MCM complex



Archaeal MCM complex

- DNA stimulated ATPase activity
- Helicase activity (3'->5') observed in archaea and eukarya
- MCM travels with the replisome
- Required for initiation and for the progression of the replication fork

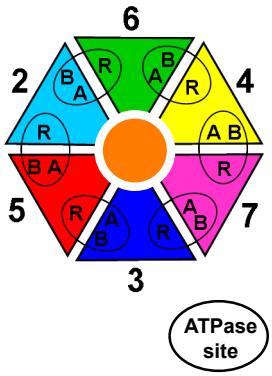
MCM helicases

MCM architecture

The AAA+ engine assembles as an **hexamer**, with the ATP binding site at the interface between subunits.

The **DNA** is expected to pass through the central hole.

ATP hydrolysis causes a change in the active site, which is transmitted and amplified so as to cause a large **conformational changes** in the subunit, which is then transmitted to the entire hexamer, causing **mechanical work** (such as "walking" along the DNA and opening up the double helix).



Helicases as combustion engines

Helicases can be depicted as an internal combustion engine with **each individual ATPase site serving as one cylinder**.

Each individual cylinder follows a defined series of events: **injection** (ATP-binding), **compression** (optimally positioning the site for hydrolysis), **combustion** (ATP hydrolysis/work generation), and **exhaust** (ADP and phosphate release).

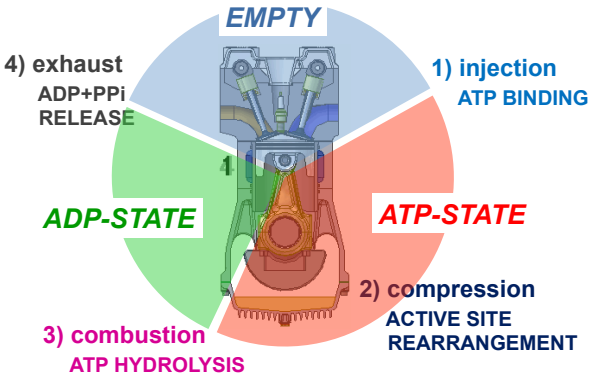
4) exhaust
ADP+PPi RELEASE

1) injection
ATP BINDING

3) combustion
ATP HYDROLYSIS

2) compression
ACTIVE SITE REARRANGEMENT

Helicases as combustion engines



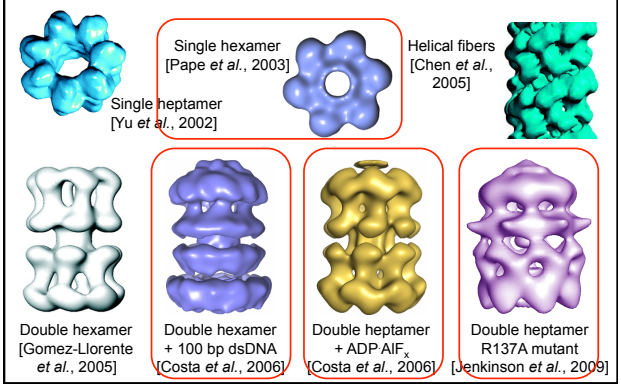
Structure & function of MCM helicases

The big questions...

- How does MCM work?
  How is the energy from ATP hydrolysis used for walking along the DNA and opening up the helix?
- How is MCM loaded onto DNA?

Structures of MCM helicases

Polymorphism as visualised by EM



Structures of MCM helicases

SsoMCM "full-length" structure

Crystal of full-length *Sso*MCM and a C-term deletion mutant were obtained.

Poorly diffracting.
Crystal structure to 4.35 A

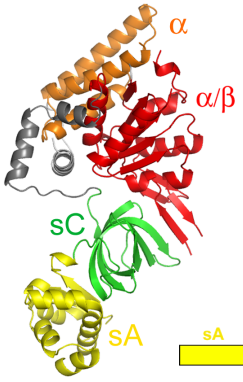
In the FL structure the C-term domain is disordered.

Only a polyalanine model was refined - but numbering indicates the register.

[Brewester *et al.*, 2008]

Structures of MCM helicases

MkaMCM2 "full-length" structure



Methanopyrus kandleri has two MCM-like genes:
- MCM1: "normal" MCM
- MCM2: lacks subdomain B, C-term, Walker A, Walker B, Sensor I, R finger.

pseudo-gene?

Crystallised as monomer.

Crystal structure to 1.9 A

[Bae *et al.*, 2009]

Structures of MCM helicases

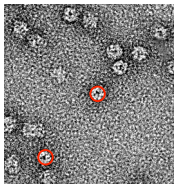
MCM+substrates

Unliganded MCM

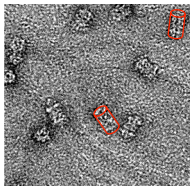
mostly single rings

MCM + DNA
MCM+DNA+ADP·AlF$_x$
MCM+DNA+AMP-PNP
MCM + ADP·AlF$_x$
MCM + AMP-PNP

mostly double rings



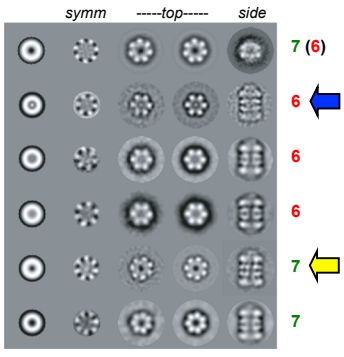
single rings, top view



double ring, side view

Structures of MCM helicases

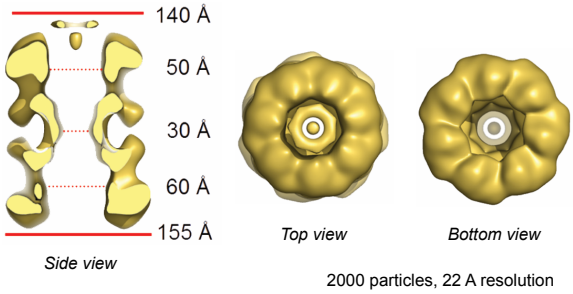
Ring stoichiometry



Structures of MCM helicases

MCM + ADP·AlF$_x$

- two heptamers head-to-head
- no electron density in the channel

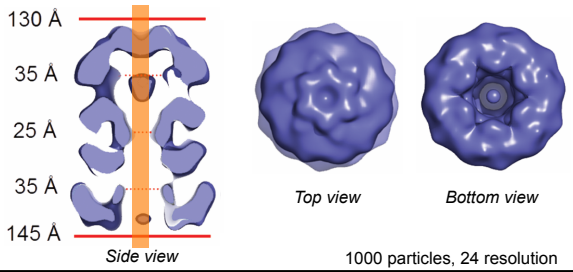


2000 particles, 22 A resolution

Structures of MCM helicases

MCM + 100 bp dsDNA

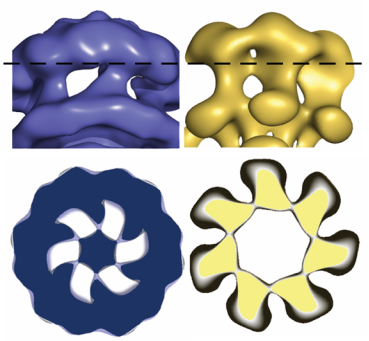
- two hexamers head-to-head
- strong asymmetry between the two hexamers
- cap closing the upper ring
- electron density in the channel



1000 particles, 24 resolution

Structures of MCM helicases

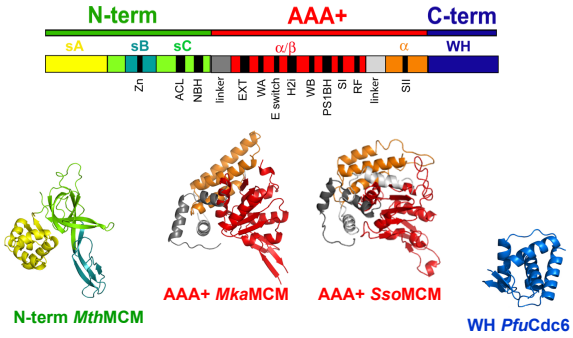
A look inside the channel



MCM + 100bp dsDNA MCM + ADP·AlF$_x$

Fitting PX models into EM maps

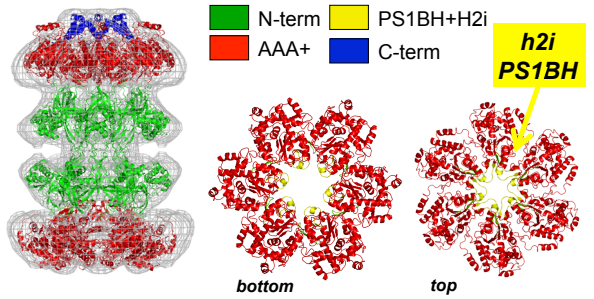
MCM crystal structures



Structure & function of MCM helicases

Conformational changes upon DNA binding

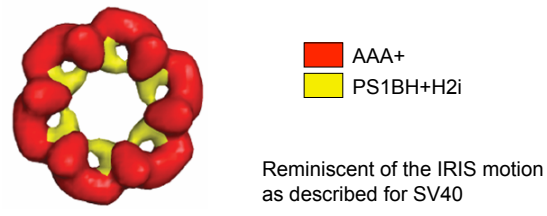
We need to envisage a conformational change to model the bottom and top AAA+ domain.



Structure & function of MCM helicases

Conformational changes upon DNA binding

If we remove the N- and C-term domains and focus only on the catalytic AAA+ domain:



Reminiscent of the IRIS motion as described for SV40

WARNING! here we are imposing six-fold symmetry...

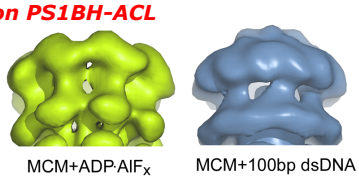
Structure & function of MCM helicases

The big questions...

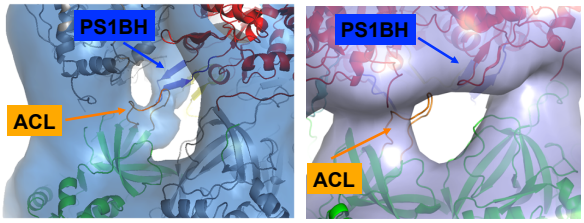
- How does MCM work?
  How is the energy from ATP hydrolysis used for walking along the DNA and opening up the helix?
- How is MCM loaded onto DNA?

Cross interactions

Interaction PS1BH-ACL

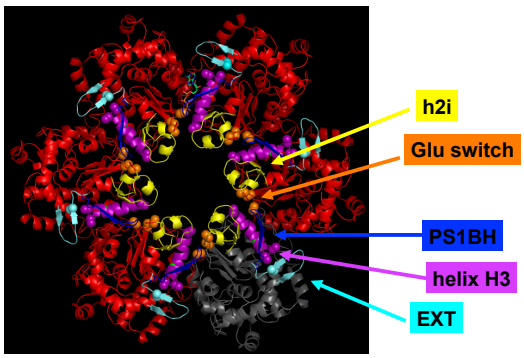


MCM+ADP·AlF$_x$ MCM+100bp dsDNA



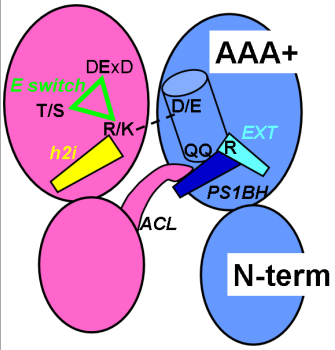
Cross interactions

The Glu switch?



Cross interactions

How to switch ON/OFF



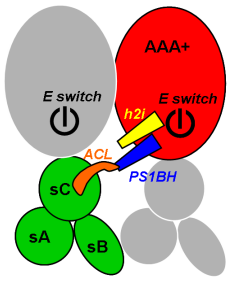
The Arg/Lys immediately precedes h2 and h2i.

The Arg/Lys can interact with a conserved Asp/Glu on an adjacent subunit, that is on helix 3, a highly conserved helix that is in contact with the PS1BH and EXT loops.

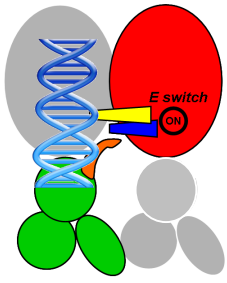
E switch "talks" to h2i (intra-subunit) and to PS1BH (inter-subunit).

Cross interactions

How to switch ON/OFF



In the **absence of DNA**: PS1BH and h2i lateral position - **E switch OFF**



In the **presence of DNA**: PS1BH and h2i facing the central channel - **E switch ON**

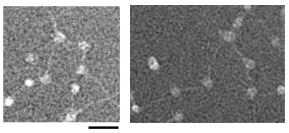
Structure & function of MCM helicases

The big questions...

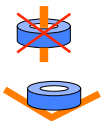
- How does MCM work?
  How is the energy from ATP hydrolysis used for walking along the DNA and opening up the helix?
- How is MCM loaded onto DNA?

CryoEM

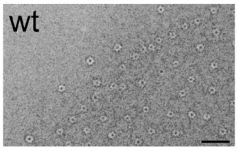
In the presence of 5kb of dsDNA



single rings,
mostly top views,
"sitting" on DNA



negatively stained micrograph (high defocus, low magnification)

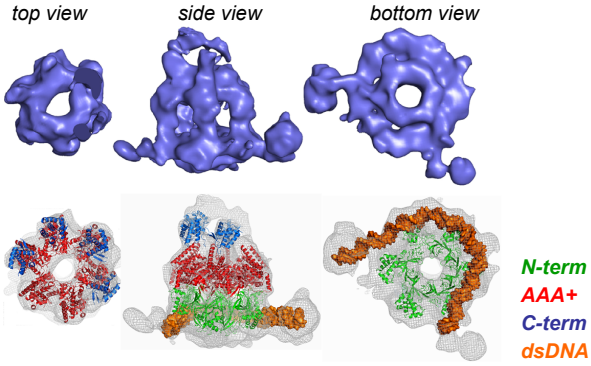


Cryo-EM micrograph

wt  hexamer

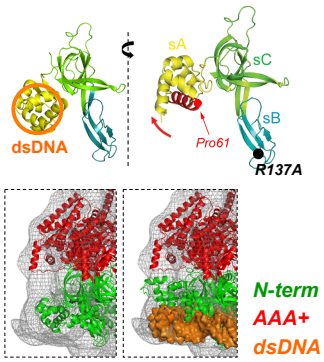
CryoEM

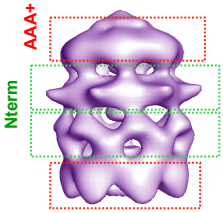
The 3D structure



CryoEM

Conformational changes of N-term domain



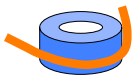


3D reconstruction of
R137A mutant:
huge swing-out of sA

Two modes of DNA binding

Role in initiation?

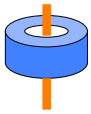
1. Initial interaction between MCM and dsDNA, with the DNA wrapping around the N-terminal domain.



Has a role in the initiation of DNA replication and assist in origin melting?

2. Canonical binding to the central channel of the MCM ring.

   The helicase is now loaded onto the DNA, ready for ATP-hydrolysis coupled translocation.



In *S. cerevisiae* two types of association between MCM2-7 and origins have been described ("associated" vs "loaded" protein)

Overview

- An introduction to single particle electron microscopy: strengths and weaknesses
- The biological problem:
  - DNA replication in eukaryotes
  - MCM helicases
  - Structural studies of MCM helicases
  - How does it work
  - How is it loaded onto DNA
- A comparison between macromolecular crystallography (MX) and single particle electron microscopy (EM)

A comparison between MX and EM

(Cryo)EM

Cryo-EM advantages

No crystals needed
Protein truly in solution (cryo! not for negative stain)
No upper size limit
Need small amount of sample
Can deal with heterogeneous samples

Cryo-EM disadvantages

Minimum size ~100-300 kDa
Not atomic resolution (10Å at best – 3Å in exceptional circumstances)
Not even "amino acid" resolution – cannot identify subunits in a complex
Methodology still being developed
Lack quality control mechanisms
Possibility of bias

BUT - very useful if combined with high-res MX or NMR structures

A comparison between MX and EM

Macromolecular crystallography

MX advantages

Atomic resolution possible
High size limit (ribosome - 2.6 MDa; viruses up to 66 MDa, 750 Å)
Rapid analysis of protein-ligand complexes (unless there are conformational changes!)

MX disadvantages

Requires large amount of protein (less with robots)
Need for crystals is a huge bottleneck
Phase problem
Crystal packing – possible artifacts?
Difficult to study conformational changes

A comparison between MX and EM

When they compete directly (such as the structure of the ribosome or RNA polymerase) MX gives far more information.

But EM can be very useful

- where it is difficult to produce enough sample for MX
- for large complexes purified directly from cells
- for large molecules or assemblies that are difficult to crystallise
- for intrinsically polymorphic samples
- to look at a number of different physiological states and observe conformational changes

Plus EM techniques are still in their "infancy" – we can expect new developments:

- in computational techniques
- in EM tomography