

Image Processing in cryoEM:

“Sunny” as ever

José-Maria CARAZO lab

Centro Nacional de Biotecnología-CSIC

Instruct Image Processing Center

carazo@cnb.csic.es



What we do...

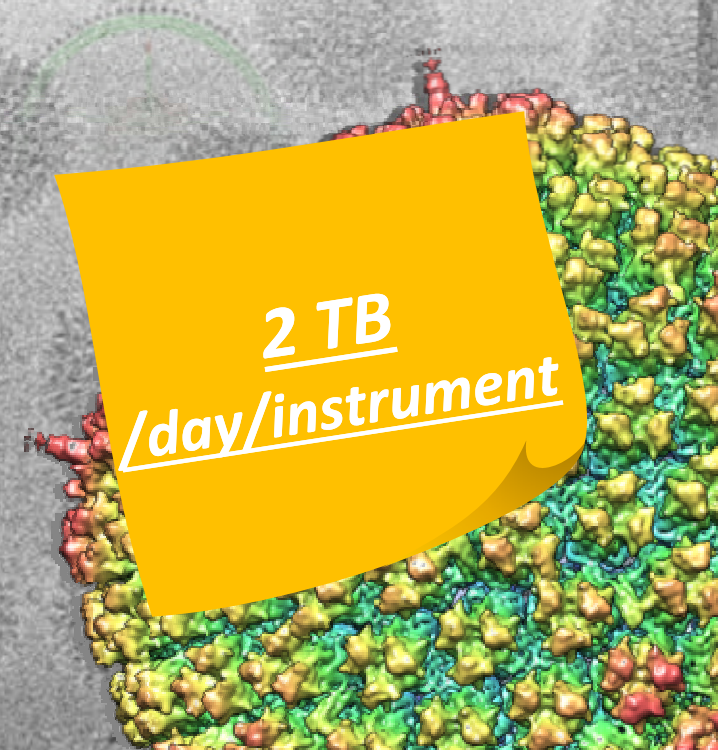
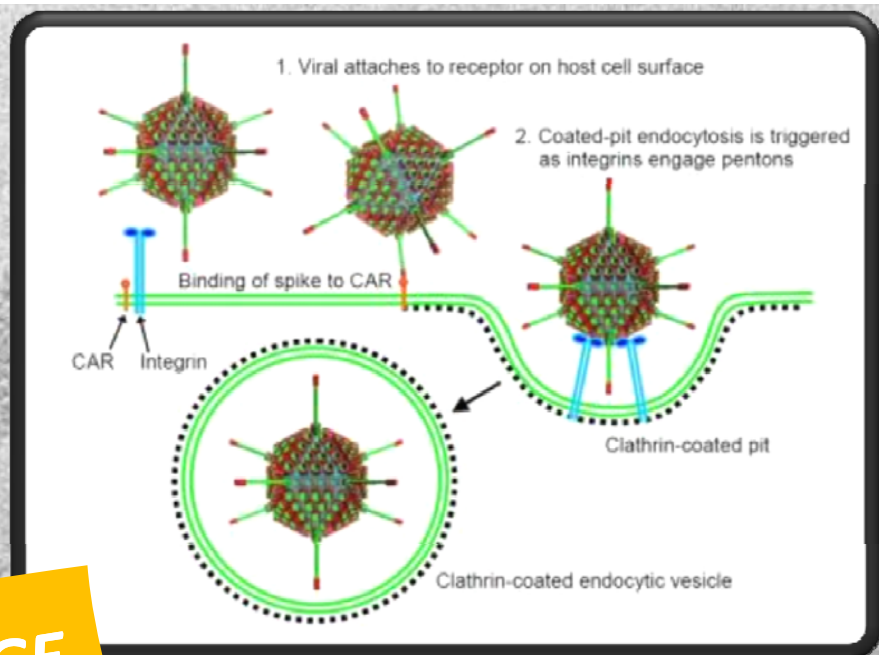
... at the INSTRUCT Image Processing Center (Madrid)



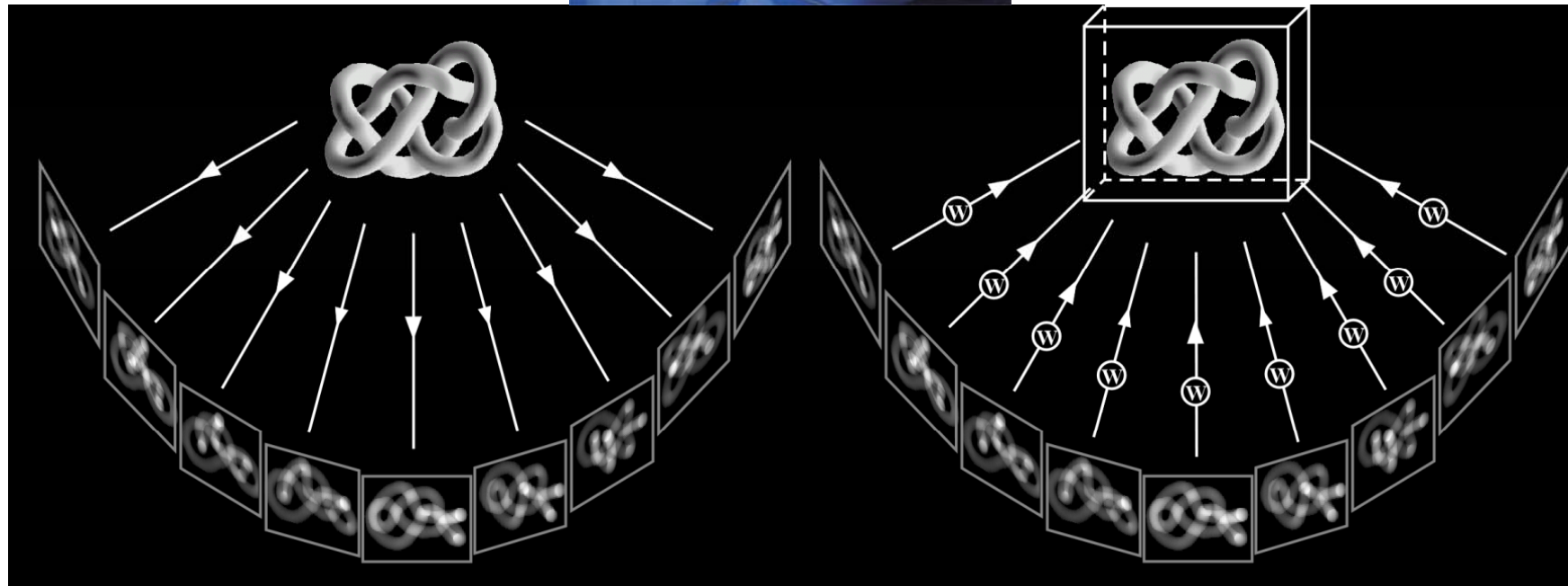
Cryo Electron
Microscopy

NEW IMAGE
PROCESSING
METHODS

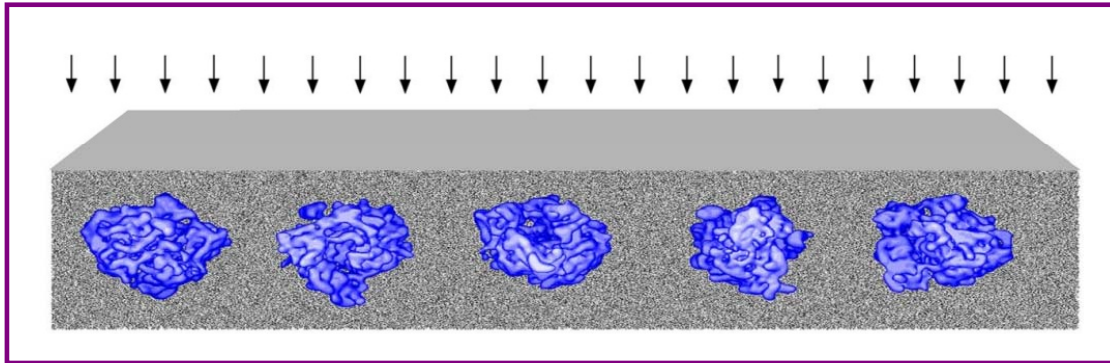
2 TB
/day/instrument



3D reconstruction Principles



The cryo EM “pledge”



Trapped in solid amorphous ice, each individual macromolecule can be imaged, revealing its internal flexibility, often linked with biological function



An example of cryo EM structure determination

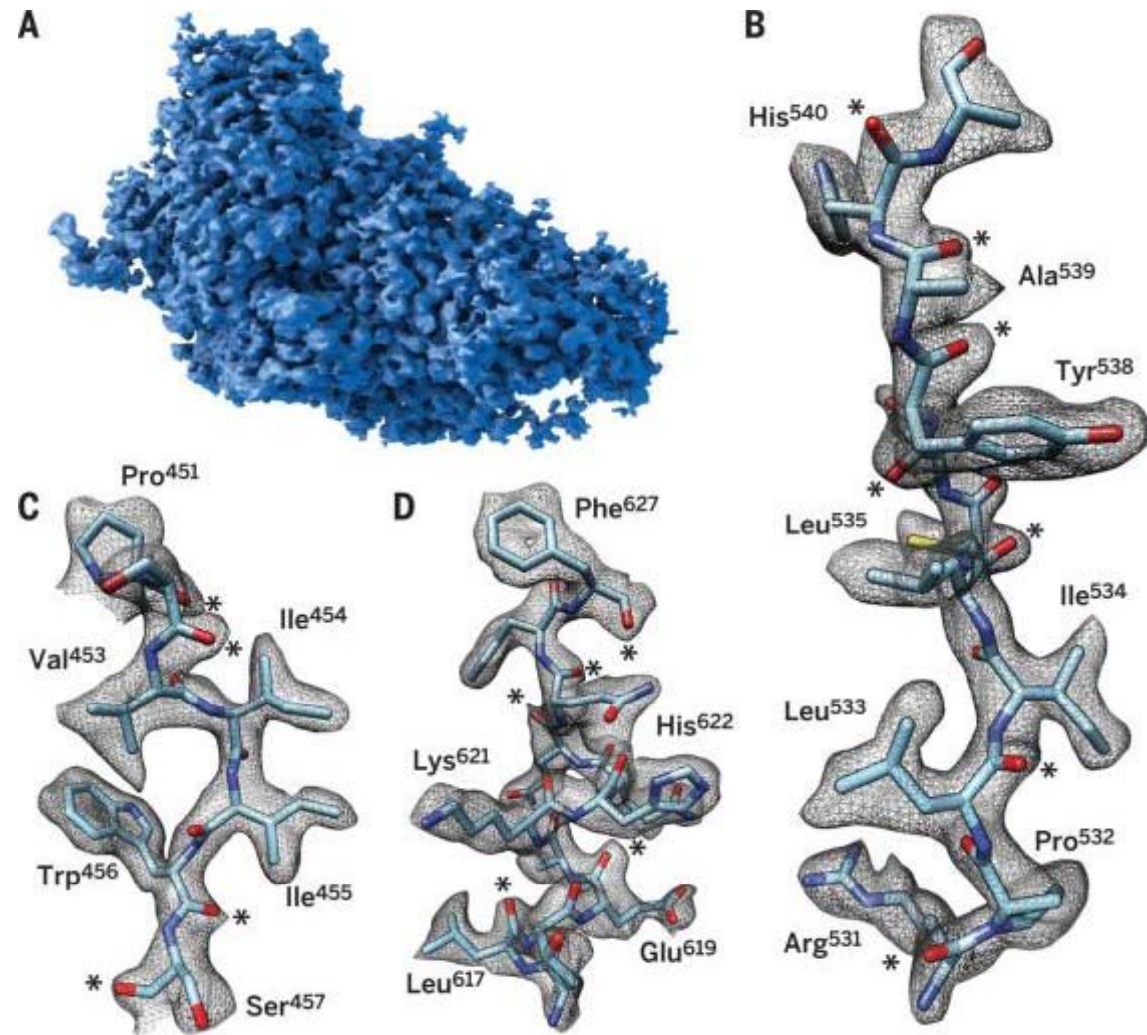
Scienceexpress

2.2 Å resolution cryo-EM structure of β -galactosidase in complex with a cell-permeant inhibitor

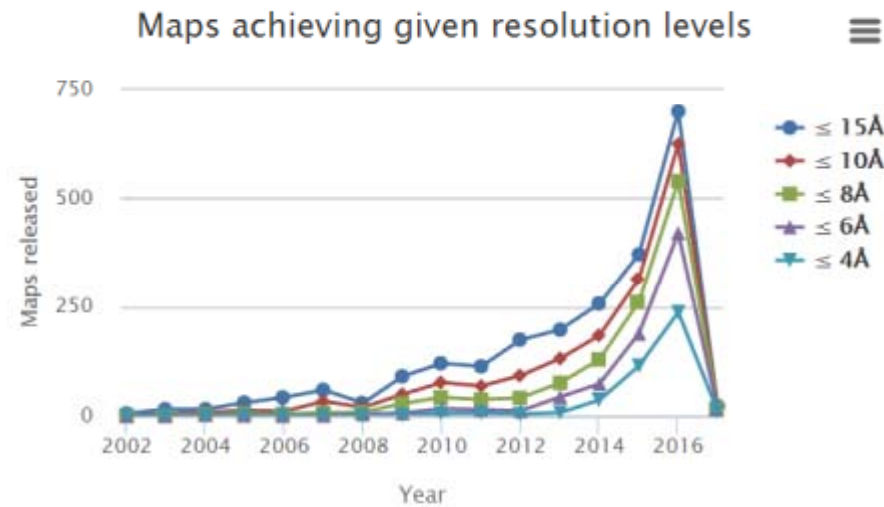
Alberto Bartesaghi,^{1*} Alan Merk,^{1*} Soojay Banerjee,¹ Doreen Matthies,¹ Xiongwu Wu,² Jacqueline L. S. Milne,¹ Sriram Subramaniam^{1†}



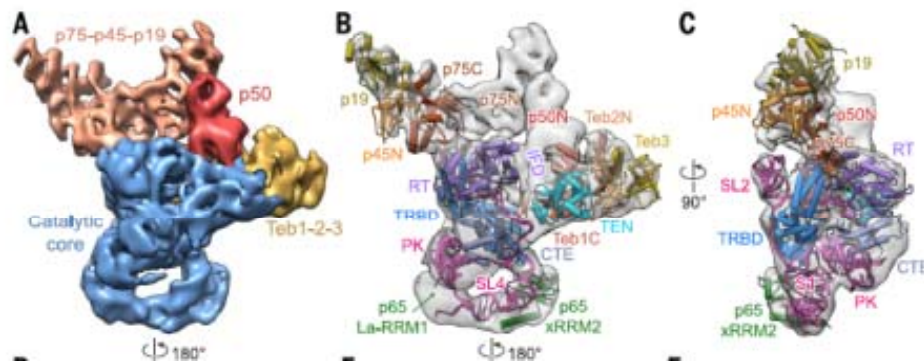
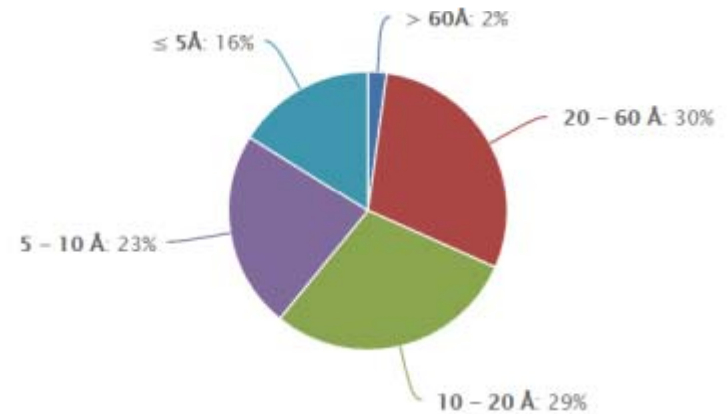
Fig. 1 Cryo-EM density map of the β -Gal-PETG complex at 2.2 Å resolution.



In 2012, FOUR structures below 4 Å were solved. Now...

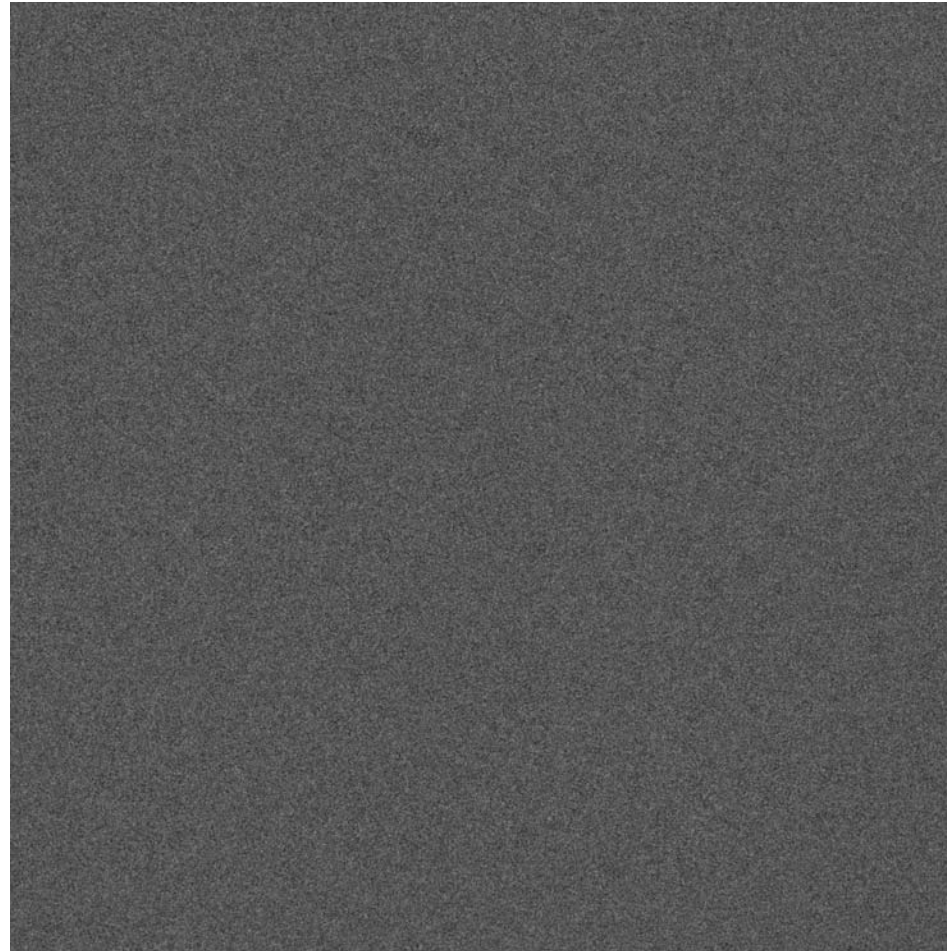


Resolution distribution for released maps



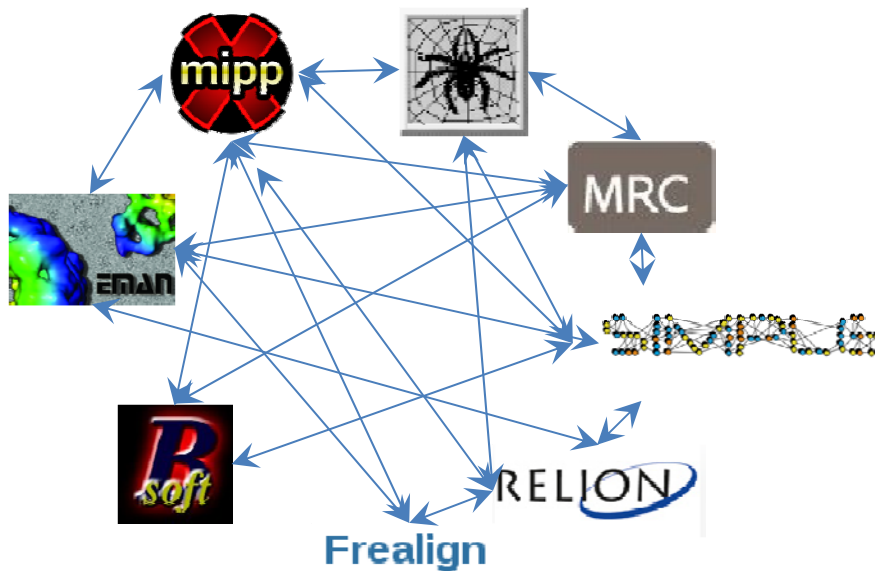
www.ebi.ac.uk/pdbe/emdb/

A typical Direct Electron Image



The EM field needs software integration

Using different EM software packages may be difficult, specially to have a clear “reproducibility track record”





SCIPION

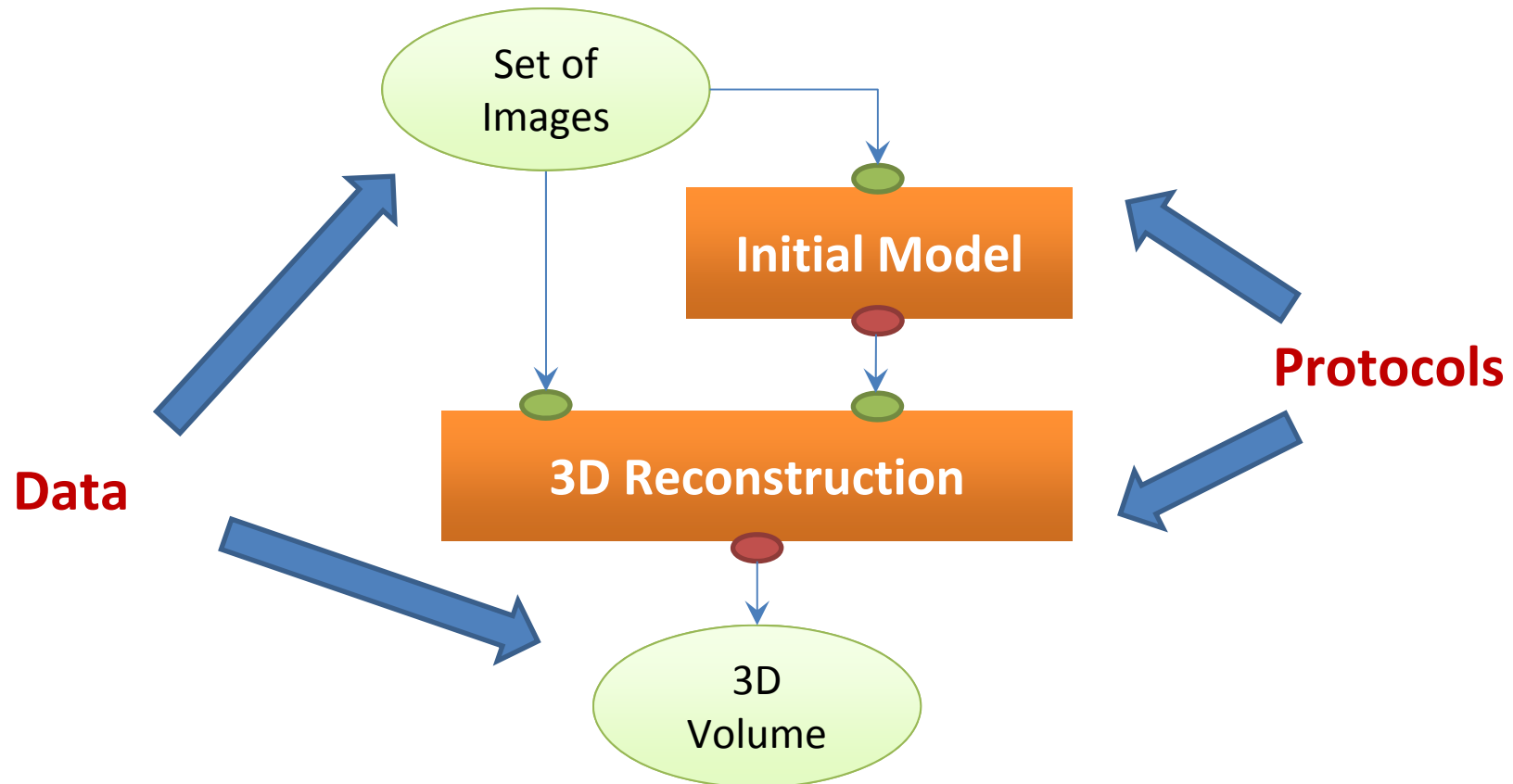
A Software Framework
for Integration, Reproducibility
and Validation in Cryo-EM



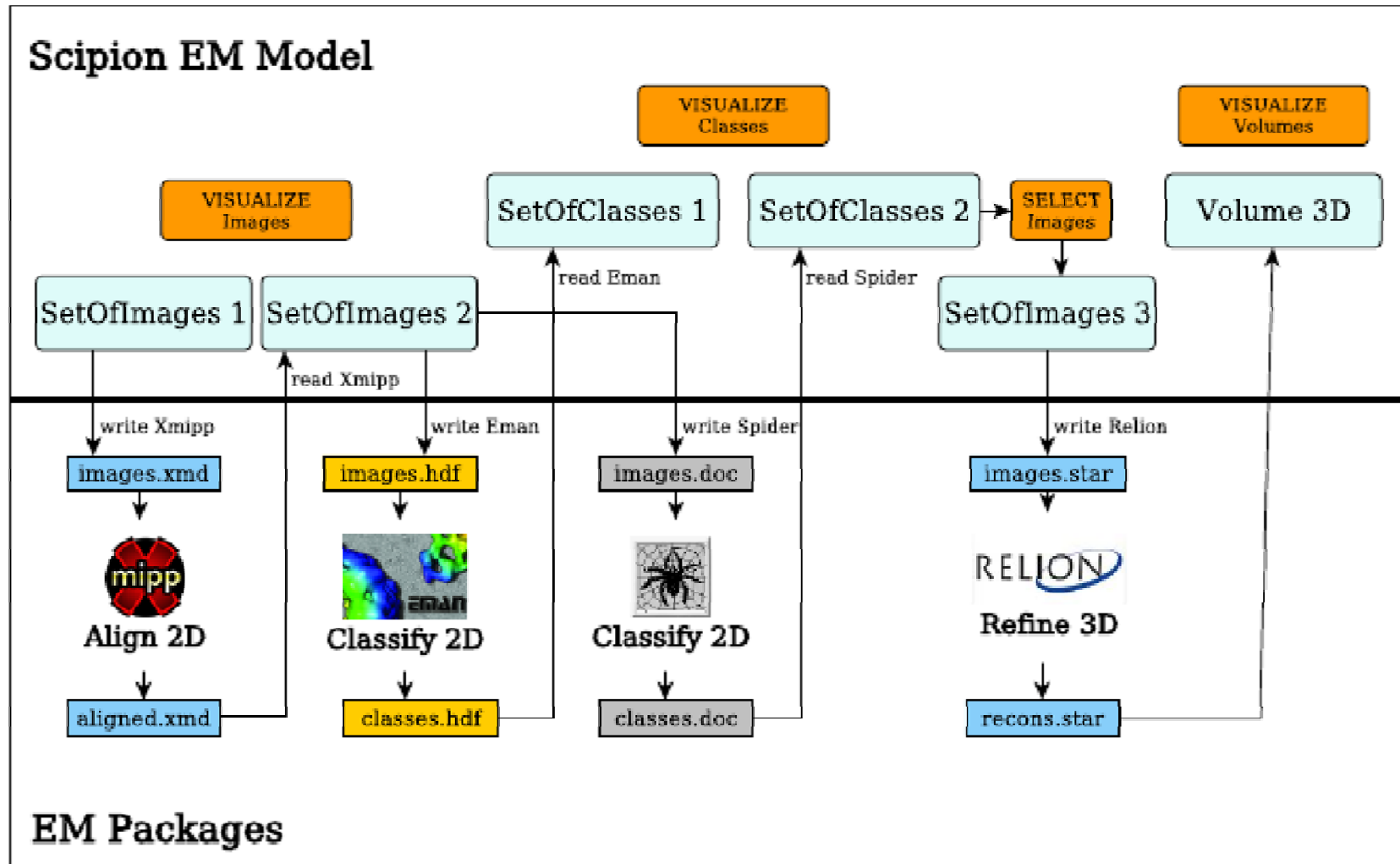
Goal 1: Integrate EM software packages to be used in the same project.



We created a model of the EM domain



Modelling but need of conversion functions for each package



Goal 2: Full project traceability,
improving reproducibility.



We track all the steps performed in a project

The screenshot displays the SCIPION software interface. On the left is a navigation pane with a tree view of protocols under 'Protocols SPA'. The main area shows a workflow tree for a 'PROJECT' with the following steps:

- PROJECT
- scipion - import particles (finished)
- scipion - filter particles (finished)
- scipion - align opsr (finished) and scipion - align pairwise (finished)
- scipion - custom mask (finished)
- scipion - copco (finished)
- scipion - classify ward (finished), scipion - classify kmeans (finished), and scipion - classify dlday (finished)

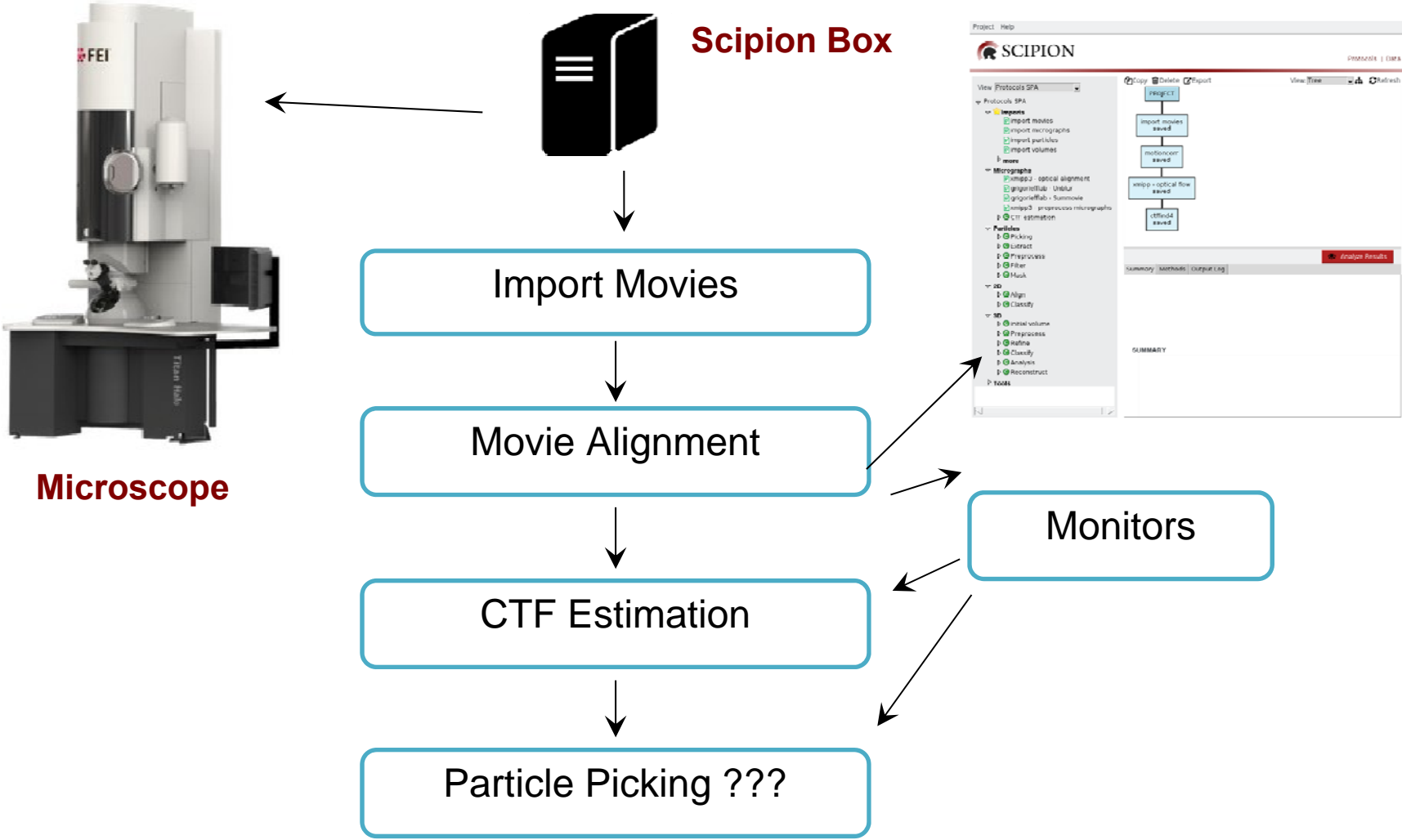
Below the tree is a summary panel with tabs for 'Summary', 'Methods', and 'Output Log'. The 'Summary' tab is active, showing input and output details for the 'scipion - align pairwise' step. The input is 'inputParticles [from scipion - filter particles -> outputParticles SetOfParticles (76 items, 100 x 100, 3.50 A/px)]' and the output is 'scipion - align pairwise -> output/Avragn Particle (100 x 100, 3.50 A/px)'. A 'SUMMARY' section indicates 'Radius range: 5 - 44'.



Goal 3: Execute complete workflows in an automated manner.



Run workflows automatically and in streaming

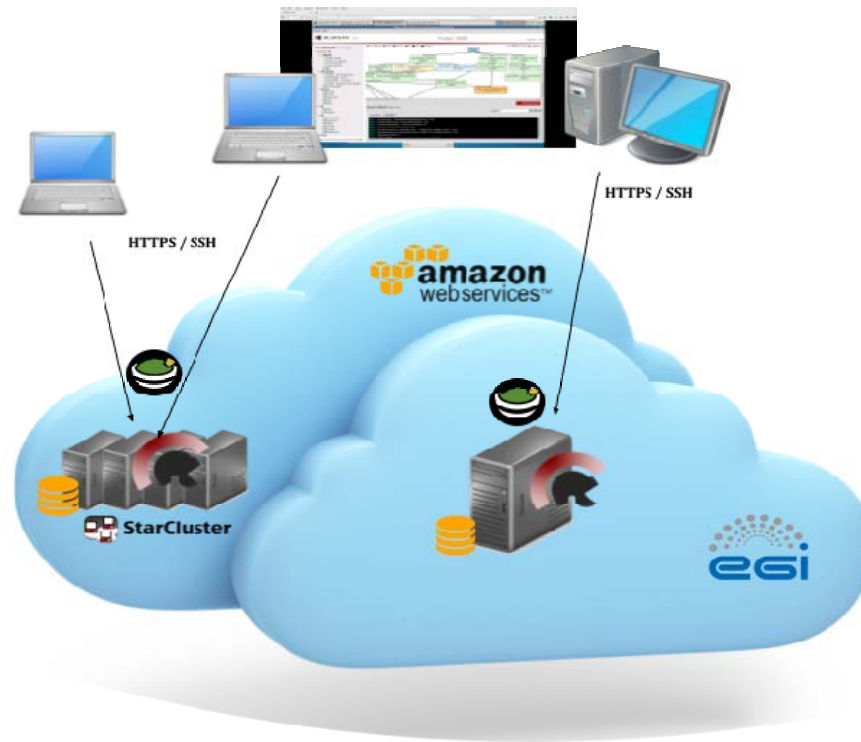


Goal 4: Local and Cloud deployment

Fully deployed on Amazon cloud and on the European Federated Cloud!



Scipion on the Cloud



West-Life

MoBrain

 Instruct
Integrating
Biology





EGi-Engage

Scipion cloud deployment for MoBrain

D6.14

Date	03 March 2017
Activity	WP6
Lead Partner	CNB – CSIC
Document Status	FINAL
Document Link	https://documents.egi.eu/document/2982

ScipionCloud images on AWS EC2 catalogue and EGI AppDB

ScipionCloud includes:

- Ubuntu 14.04

- Scipion 1.0.1

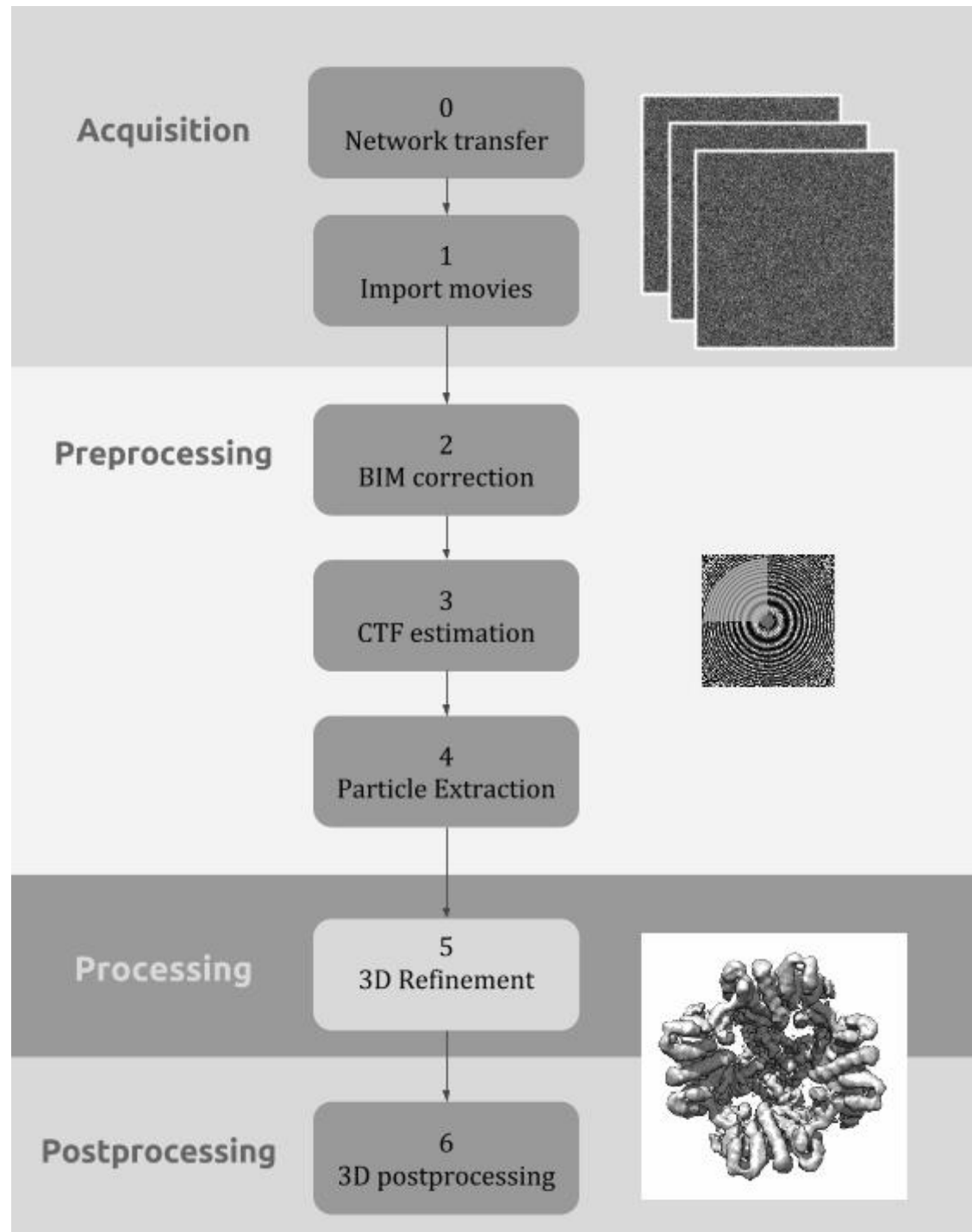
- Remote desktop solution Guacamole

- Starcluster tool for cluster management (only for AWS EC2)

- Nvidia CUDA for GPU support (currently only for AWS EC2)

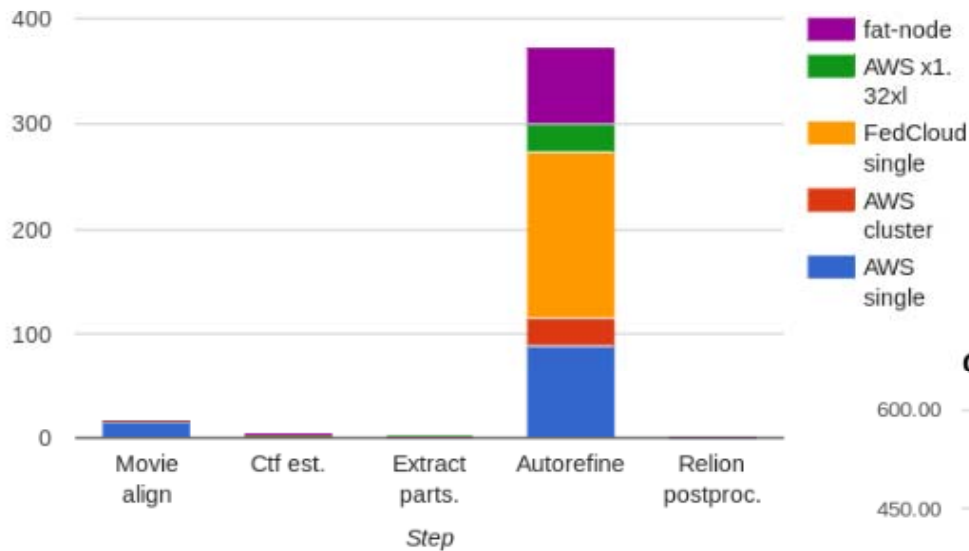
- Preinstalled EM packages: Xmipp, Relion, Ctffind, MotionCorr, Spider, EMAN, Bsoft, Chimera, Frealign, Resmap, Summovie

Documentation on <https://github.com/I2PC/scipion/wiki/ScipionCloud>



From	To	Protocol	Avg Speed (MB/s)	Max Speed (MB/s)	Cost(\$)
CNB (Spain)	AWS Ireland	rsync+ssh	30	30	10 TB disk: 32\$. Instance: 15\$. Total: 47\$
		bbcp	75	87	
EBI (UK)	AWS Ireland	Aspera FASP	75	110	
AWS Ireland	AWS Ireland	rsync+ssh	30	30	
CNB (Spain)	CESNET FedCloud	rsync+ssh	37	37	

Profiling summary (elapsed time/step, hours)



Cost distribution (\$)



Scipion Cloud in production

- Most courses now use Scipion Cloud

NeCEN

[ABOUT NECEN](#)

[SERVICES](#)

[USER ACCESS](#)

[MICROSCOPES](#)

[PEOPLE](#)

[RESEARCH PROJECTS](#)

CRYO-EM SCHOOL

- Most Pharma work made on Cloud
- Ideal to leverage computer usage

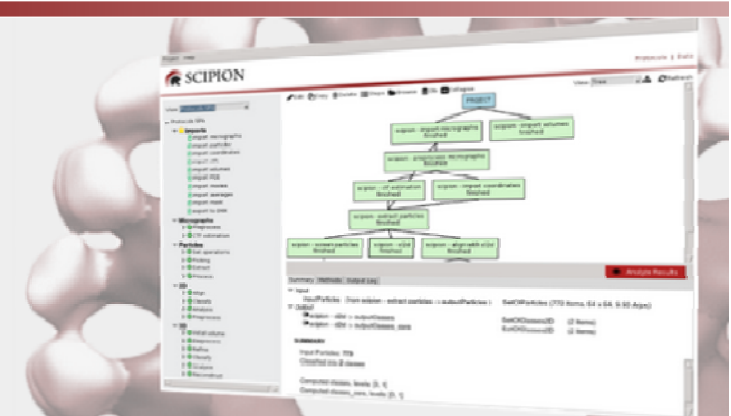
Scipion is open-source and freely available at <http://scipion.cnb.csic.es>



Integration, Reproducibility and Analysis

Scipion is an image processing framework to obtain 3D models of macromolecular complexes using Electron Microscopy.

Download



Documentation

Find out more information about Scipion for both users and developers. Check how to install Scipion, the list of integrated protocols and some introductory tutorials. Learn how to extend Scipion with new protocols, its API or description of our development tools.

Web Tools

Give a try to our online processing workflows, which are a subset of the protocols in Scipion to provide a first try without any local installation. Current web-tools include initial volume estimation, movie alignment and local resolution (ResMap).

News and Events

Keep an eye on latest Scipion news. Check out about bugfixes, new features and release plan. Don't miss the next Scipion workshop or any related event. Don't hesitate to contact us for any feedback or if you want to organize a workshop at your institution.

© 2015 - Scipion is licensed under the GNU General Public License



Scipion web tools



In this context: What are “Web tools”

From our site (<http://scipion.cnb.csic.es/>):

“online processing workflows, which are a subset of the protocols in Scipion to provide a first try without any local installation.”



Web tools home page



Scipion Web Tools - **W**est-**L**ife

Single particle analysis tools

Align your movies

My movie alignment

Create your initial volume

My first map

Analyze your maps

My resolution map

Reliability tools

Explore protein interactions



Check the reliability of your data




My reliability tool







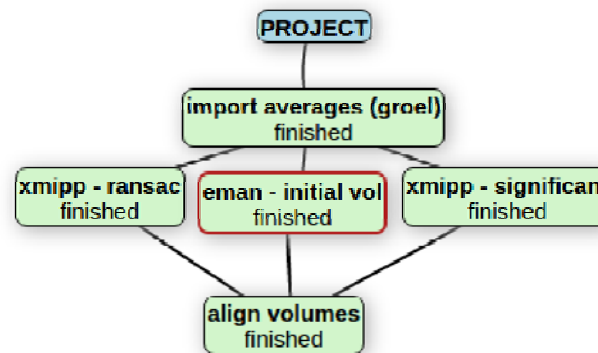
Example of a web tool

Workflow View: Initial_Volume ▾

- 1. Upload data
 - ★ import averages
- 2. Create a 3D volume
 - 📄 xmipp3 - ransac
 - 📄 eman2 - initial volume
 - 📄 xmipp3 - significant
- 3. Align volumes.
 - 📄 xmipp3 - align volumes

 Edit  Copy  Delete

 Small Tree  List  Refresh  Help



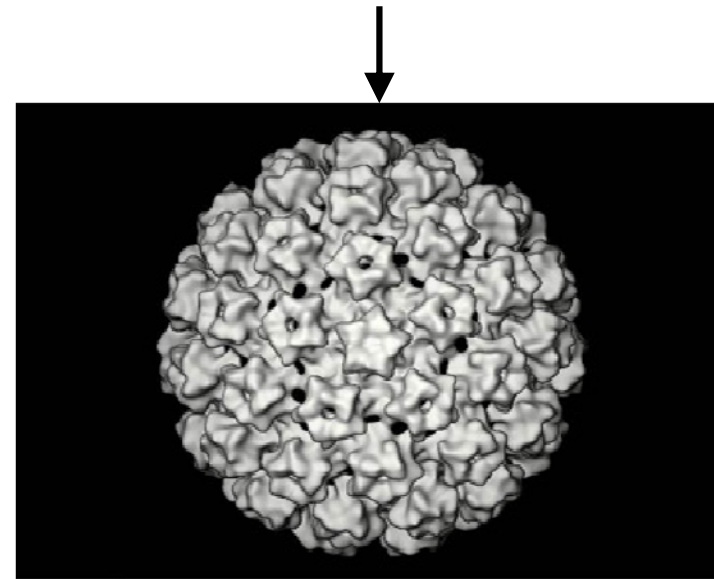
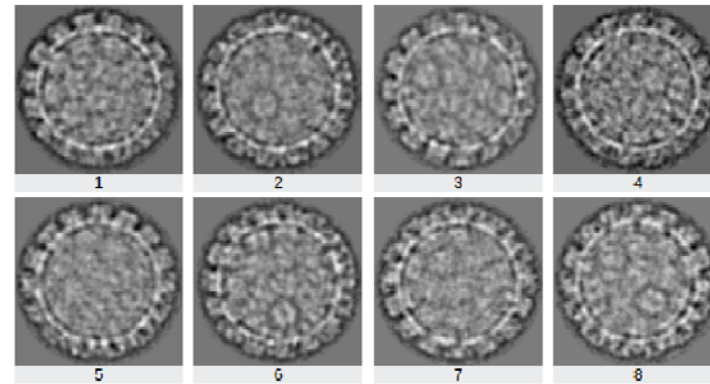
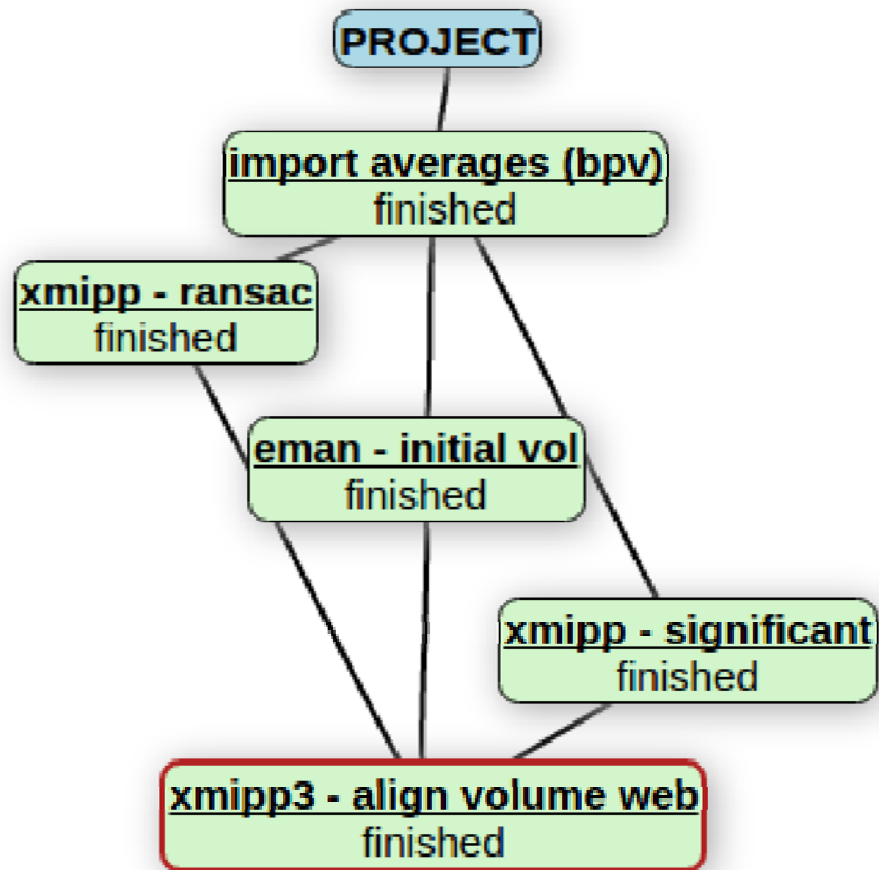
Summary | Methods | Output Logs | **Analyze Results** | Download Results

Input
➔ inputSet (from 103.inputSet.116)  SetOfAverages (31 items, 64 x 64, 1.00 A/px)

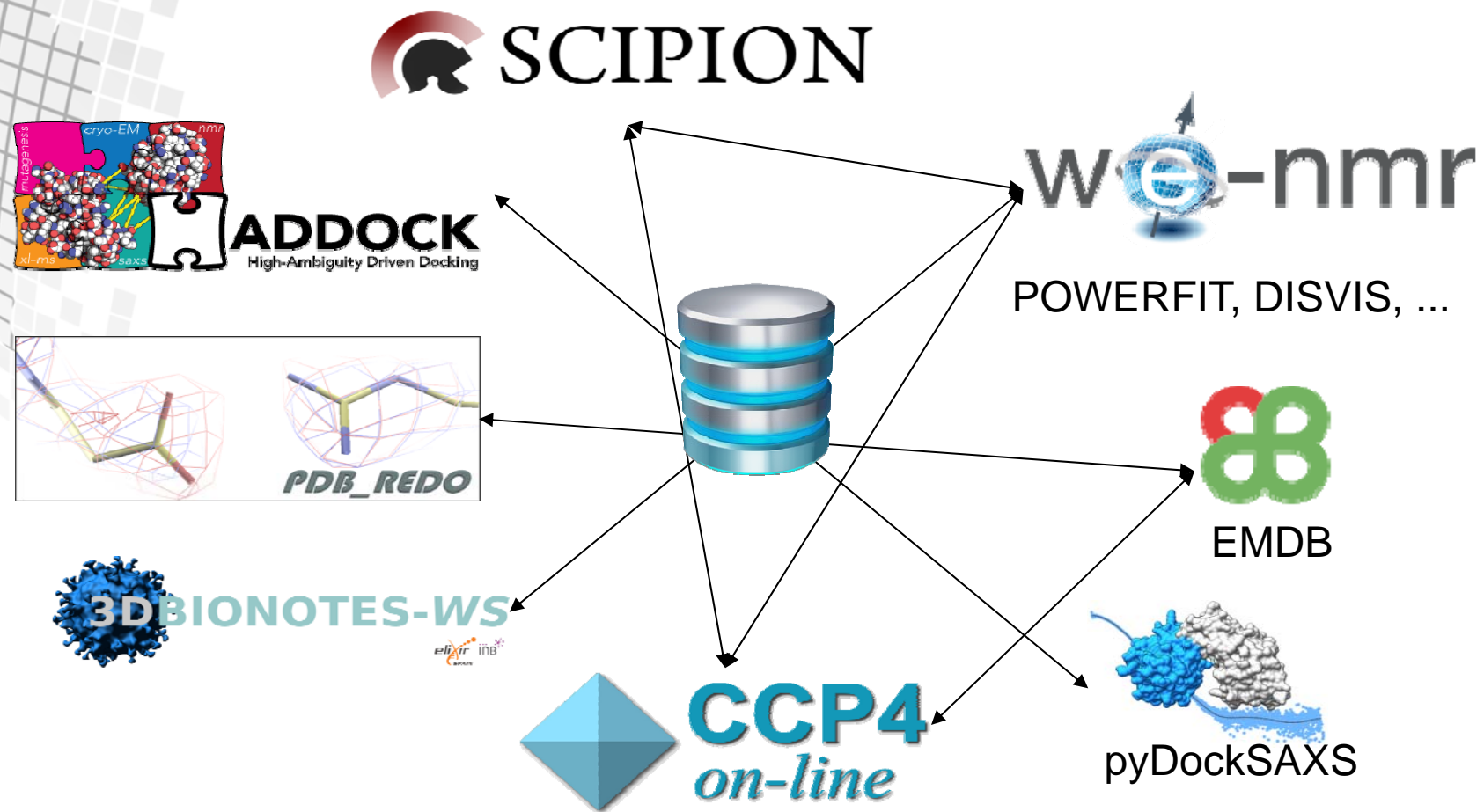
Output
➔ outputVolumes  SetOfVolumes (10 items, 64 x 64 x 64, 1.00 A/px)

Summary
Input Images: 2.outputAverages.47
Output initials volumes: None

My first map: Create your initial volume



Interconnecting Structural Biology resources: West-Life



And the people behind.....

- Carlos Oscar Sanchez-Sorzano and Roberto Marabini
- Javier Vargas
- Joan Segura
- Jose Luis Vilas and Ruben Sanchez
- Pablo Conesa, Jose Miguel de la Rosa and Scipion team
- And all the talented predocs, postdocs and engineers along the years!



And the people to come...

Always looking for talented
engineers, predocs and
postdoc fellows!