Afonso M.S. Duarte

ITQB-UNL, Portugal

Laboratory of Biological Energy Transduction

(http://www.itqb.unl.pt/labs/biological-energy-transduction)

EGI Champion

(http://www.egi.eu/community/egi_champions/Afonso_Duarte.html)

INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
/UNL

Knowledge Creation

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, INOVAÇÃO E DO ENSINO SUPERIOR

MARIE CURIE ACTIONS

egi

Structural Biology in the Grid

a case study

The importance of molecular modeling applications

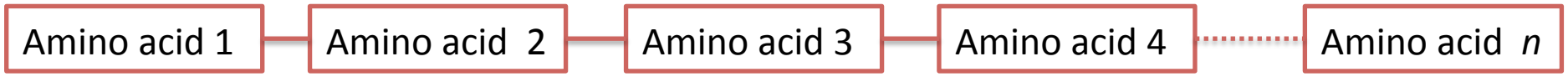
The role of the Grid

Some examples

Proteins are biomolecules that play key functions in living cells

Proteins consist of one or more chains of amino acid residues connected in a linear fashion.

Chain 1

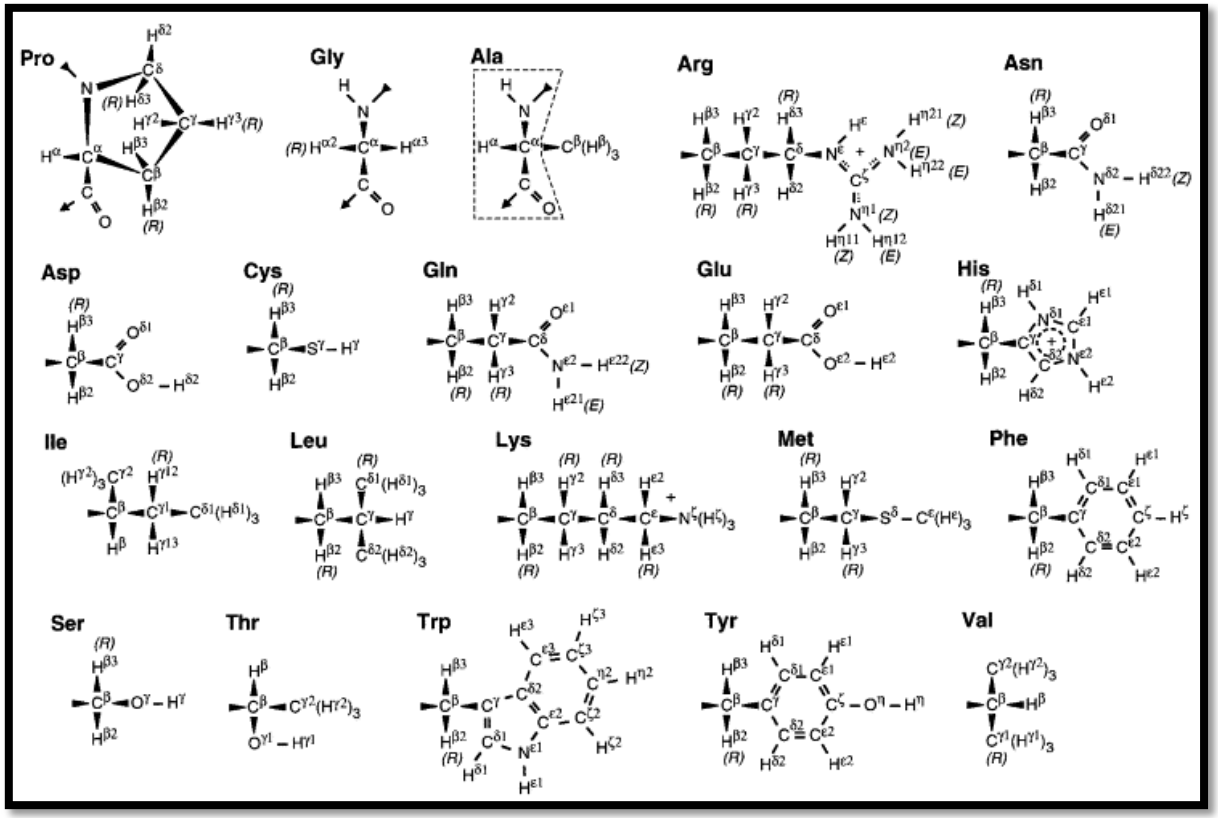


These chains can then fold and form different 3D structures. Proteins perform an array of functions within living organisms, and are center to almost all processes in life

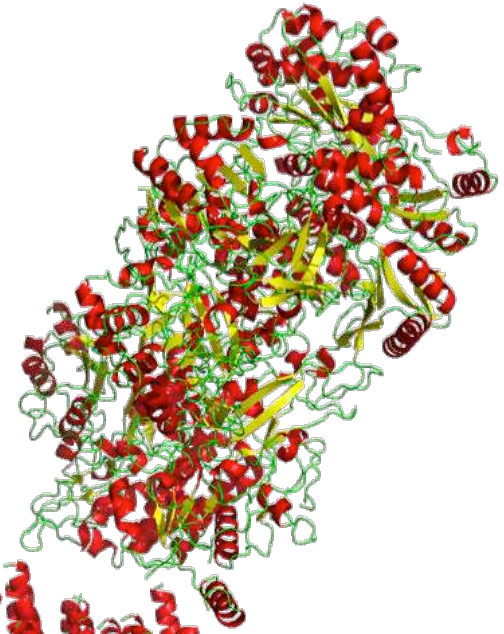
In Nature there are 20 different amino acids that buld-up proteins (http://www.bmrb.wisc.edu/referenc/commonaa.php)

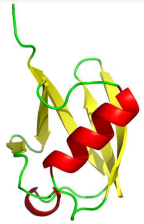
These have different chemical properties that allow the build up of proteins with different 3D structures and functions in cells

20 amino acids can be combined into proteins with different sizes, shapes and functions

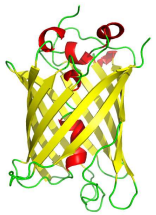


←20 ≠ amino acids

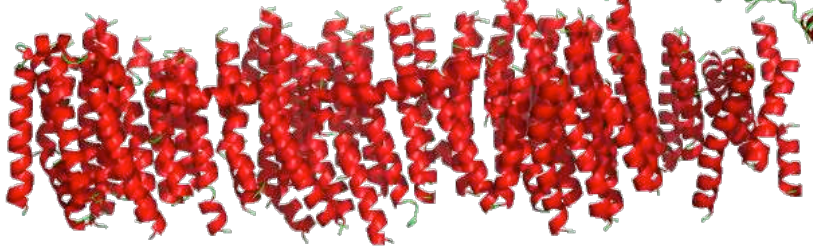




Ubiquitin



GFP



Complex I from the respiratory chain

20 Å

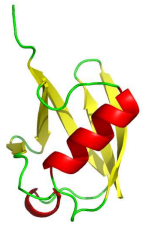
John L Markley, Ad Bax, Yoji Arata, C.W Hilbers, Robert Kaptein, Brian D Sykes, Peter E Wright, Kurt WüthrichRecommendations for the presentation of NMR structures of proteins and nucleic acidsJournal of Molecular Biology, Volume 280, Issue 5, 31 July 1998, Pages 933–952

Protein Structure

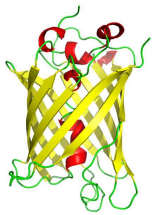
- Protein structures are determined by techniques that provide atomic resolution (i.e. where atoms are individually detected and a three-dimensional model can be constructed - resolution below 4 Å)

  – X-Ray Crystallography

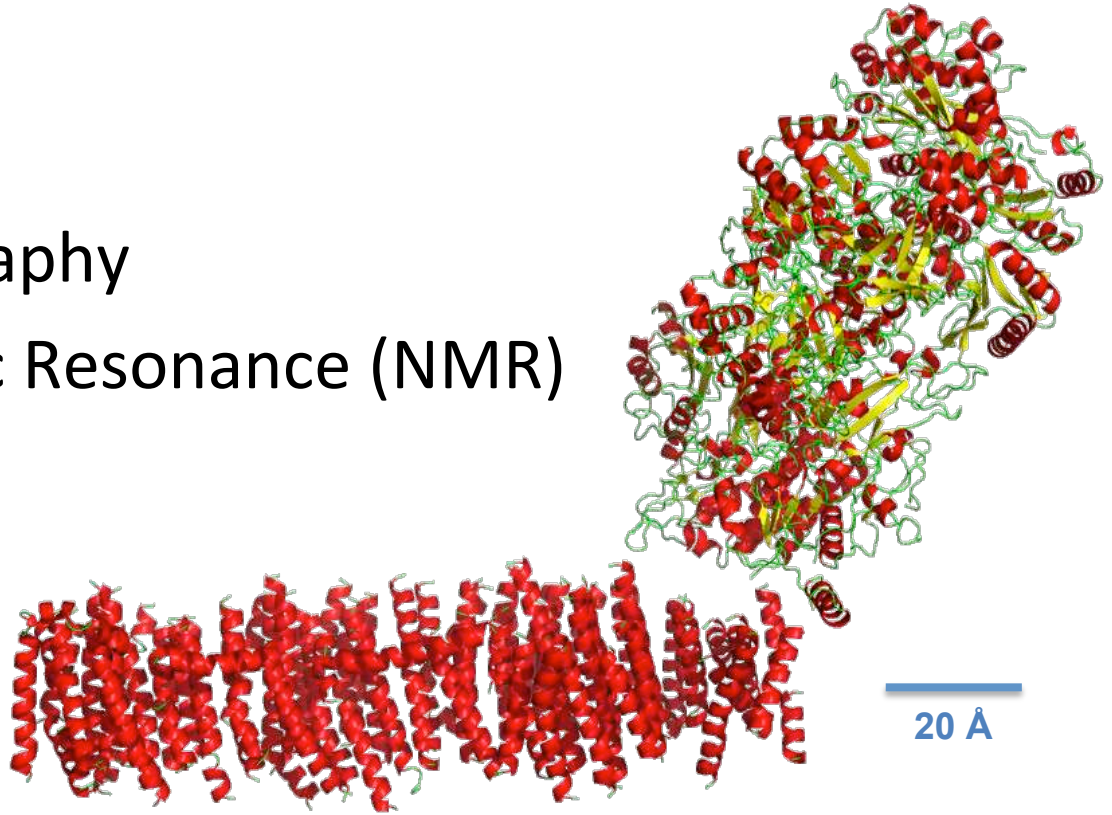
  – Nuclear Magnetic Resonance (NMR) Spectroscopy



Ubiquitin



GFP



Complex I from the respiratory chain

John L Markley, Ad Bax, Yoji Arata, C.W Hilbers, Robert Kaptein, Brian D Sykes, Peter E Wright, Kurt WüthrichRecommendations for the presentation of NMR structures of proteins and nucleic acidsJournal of Molecular Biology, Volume 280, Issue 5, 31 July 1998, Pages 933–952

The PDB format

Record type

Number in the pdb file

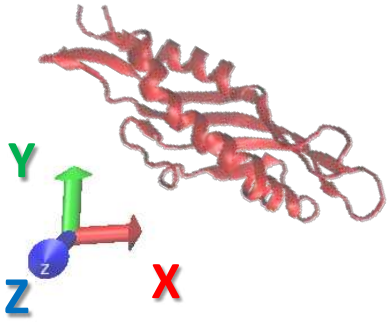
Type of atom

Amino acid

Chain name

Number in the protein sequence

Other molecular properties




```
ATOM      1  N   TRP B  11      17.076  76.256  33.096  1.00 92.88           N
ATOM      2  CA  TRP B  11      15.670  76.465  33.549  1.00 92.65           C
ATOM      3  C   TRP B  11      15.[illegible]  1.00 91.65           C
ATOM      4  O   TRP B  11      14.610  77.349  35.529  1.00 91.17           O
ATOM      5  CB  TRP B  11      15.[illegible]  1.00 93.74           C
ATOM      6  CG  TRP B  11      13.[illegible]  1.00 95.26           C
ATOM      7  CD1 TRP B  11      12.778  74.223  34.753  1.00 95.40           C
ATOM      8  CD2 TRP B  11      12.608  76.144  33.606  1.00 95.71           C
ATOM      9  CE2 TRP B  11      11.319  75.758  34.036  1.00 96.10           C
ATOM     10  CE3 TRP B  11      12.747  77.326  32.862  1.00 95.79           C
ATOM     11  NE1 TRP B  11      11.451  74.585  34.733  1.00 95.92           N
ATOM     12  CZ2 TRP B  11      10.175  76.516  33.749  1.00 96.64           C
ATOM     13  CZ3 TRP B  11      11.610  78.079  32.577  1.00 95.43           C
ATOM     14  CH2 TRP B  11      10.342  77.669  33.019  1.00 96.33           C
ATOM     15  N   VAL B  12      16.651  78.158  35.028  1.00 90.37           N
```

Coordinates

X Z Y

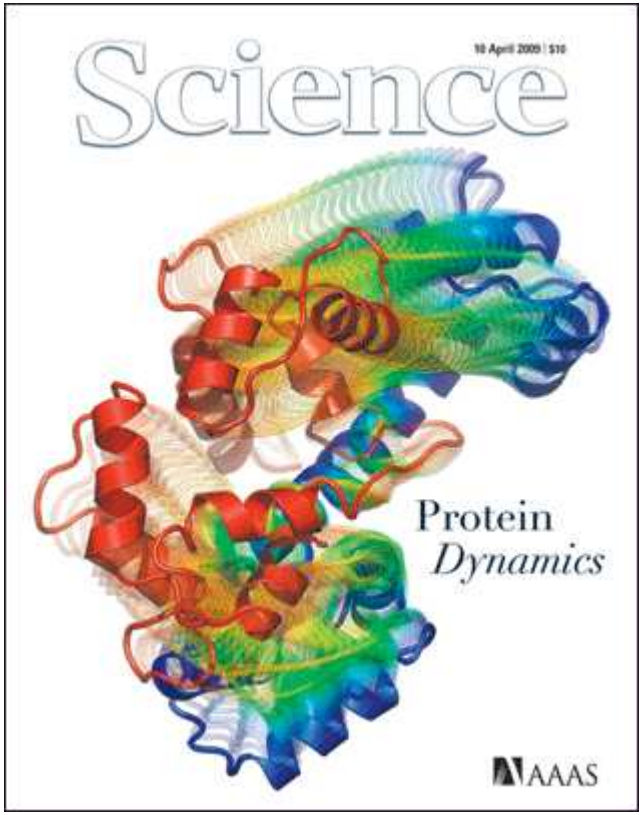
Protein dynamics

"what allows proteins to work"

Dynamics are key for protein function

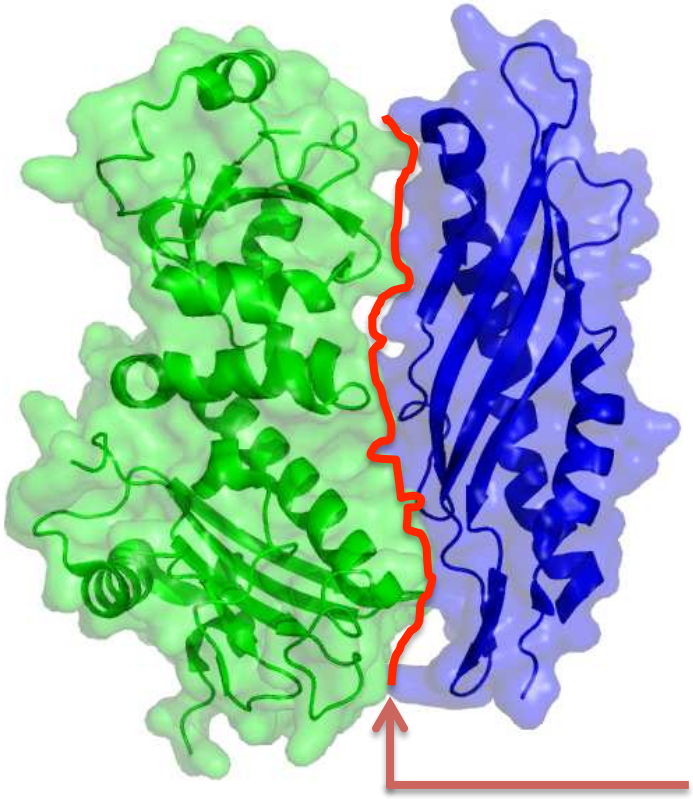
They allow structural changes fundamental for :

- Protein-protein interactions
- Protein-drug interaction
- Activation of enzymatic activity
- …



Protein – Protein interactions

"The way proteins comunicate between them"



Binding interface (i.e. binding hotspot)

Key for protein regulation in cells:

- Activation
- Inhibition
- Intermediary step in cellular processes
- …

Interaction between **Hsp90** and **Aha1** proteins.
Molecular machine involved in the cell quality control process (**Hsp90**) and its regulator (**Aha1**)

Experimental techniques

NMR spectroscopy and X-Ray crystallography provide atomic level resolution information on the different steps of protein dynamics and protein-protein interactions

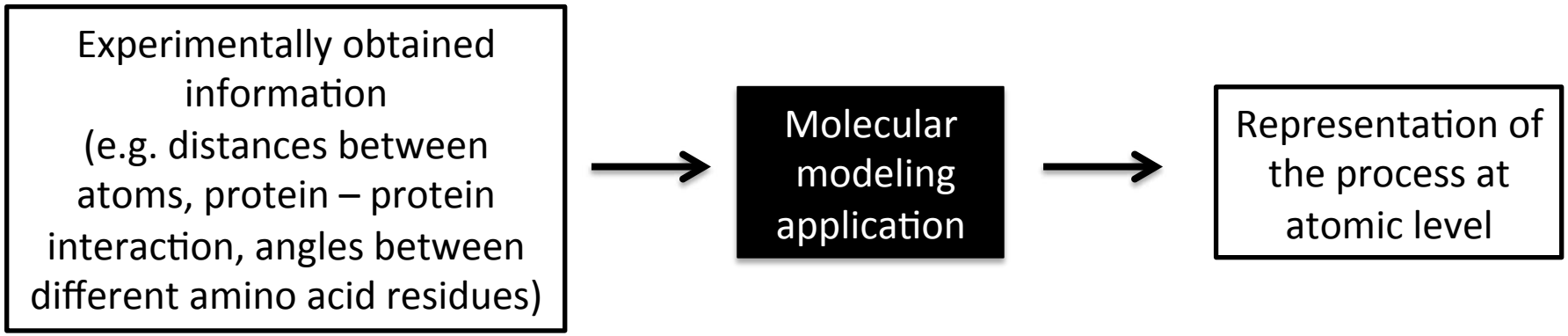
From these one obtains:

Structural constrains (e.g. distances, angles)

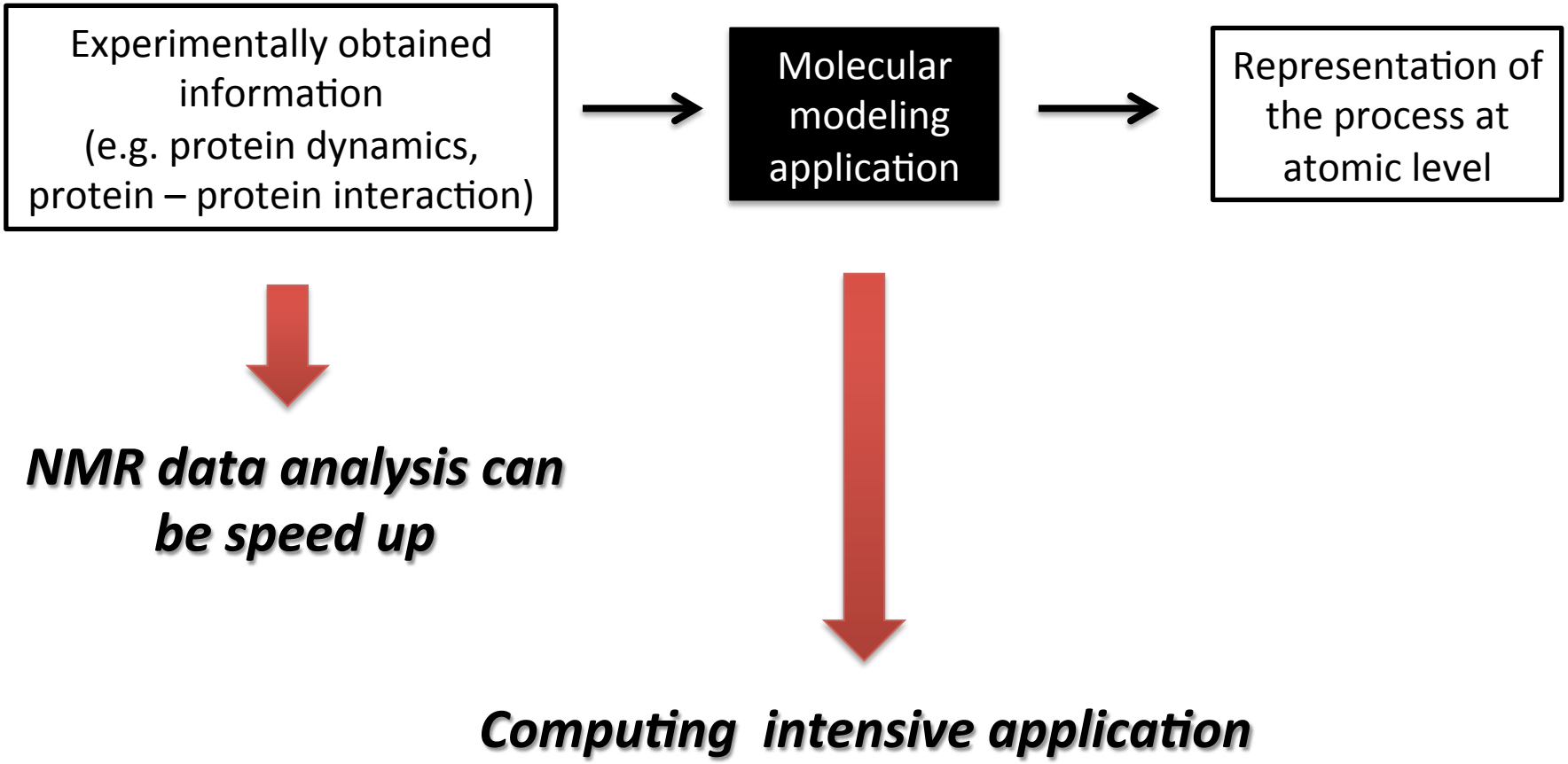
Snapshots of different steps of the process

Computational methods

To obtain a picture of the process (protein dynamics and protein-protein interactions) one uses these constrains as input for the computation of protein models that represent the process under study



Why Grid Computing?



The Virtual Organization WeNMR

"WeNMR is currently the largest Virtual Organization (VO) in life sciences. With its large and worldwide user community, WeNMR has become the first Virtual Research Community officially recognized by the European Grid Infrastructure (EGI)."

http://www.wenmr.eu/wenmr/

Tsjerk A. Wassenaar *et. al.* WeNMR: Structural Biology on the Grid **Journal of Grid Computing** 2012, 10, 4, 743-767

The Virtual Organization WeNMR

"The WeNMR project is a European Union funded international effort to **streamline** and **automate** analysis of Nuclear Magnetic Resonance (NMR) and Small Angle X-Ray scattering (SAXS) imaging data for atomic and near-atomic resolution molecular structures. "

"To facilitate the use of NMR spectroscopy and SAXS in life sciences the WeNMR consortium has **established standard computational workflows and services** through **easy-to-use web interfaces**, while still retaining sufficient flexibility to handle more specific requests. "

Tsjerk A. Wassenaar *et. al.* WeNMR: Structural Biology on the Grid **Journal of Grid Computing** 2012, 10, 4, 743-767

The Virtual Organization WeNMR

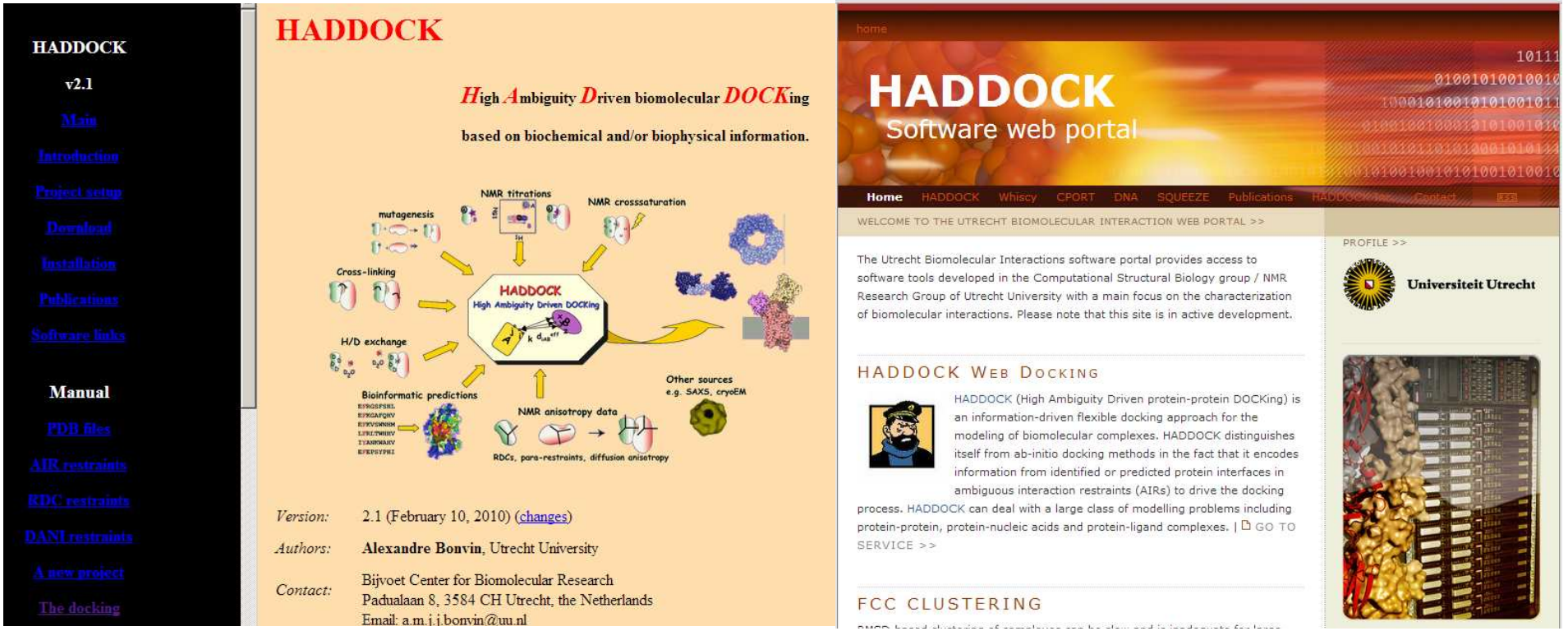
Gaining access to the Grid and to the VO

http://www.wenmr.eu/wenmr/

HADDOCK

High Ambiguity Driven biomolecular DOCKing

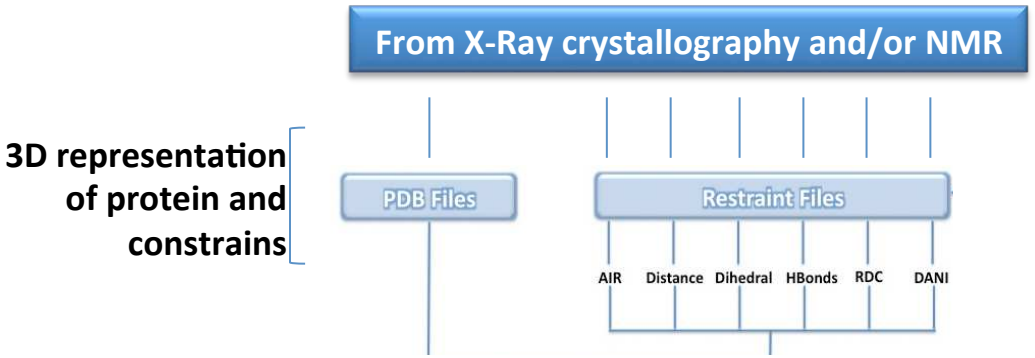
based on biochemical and/or biophysical information.



http://www.nmr.chem.uu.nl/haddock/ http://haddock.science.uu.nl/

Protein – Protein interactions

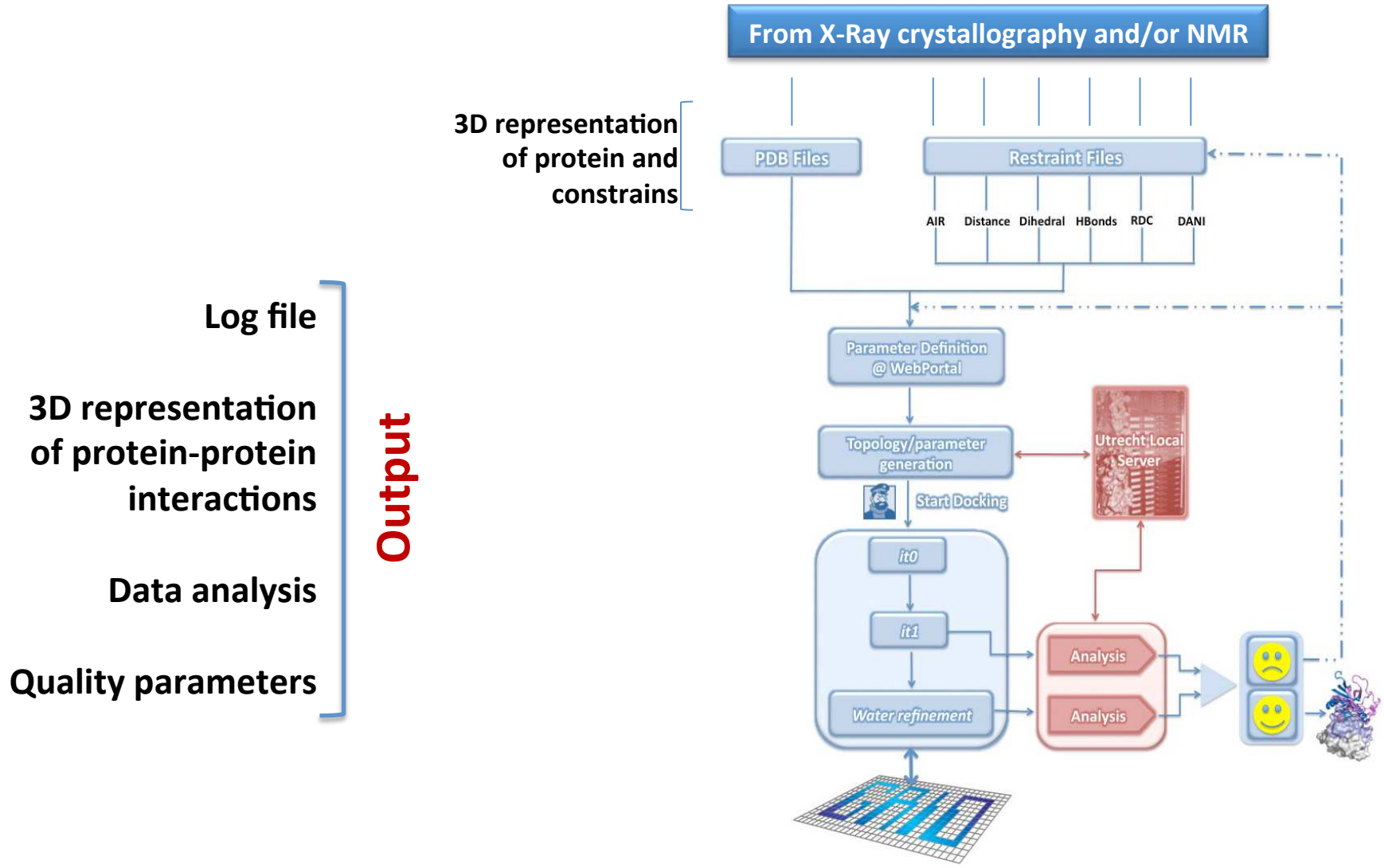
HADDOCK in the Grid



http://haddock.chem.uu.nl/enmr/haddock-tutorial.php

Protein – Protein interactions

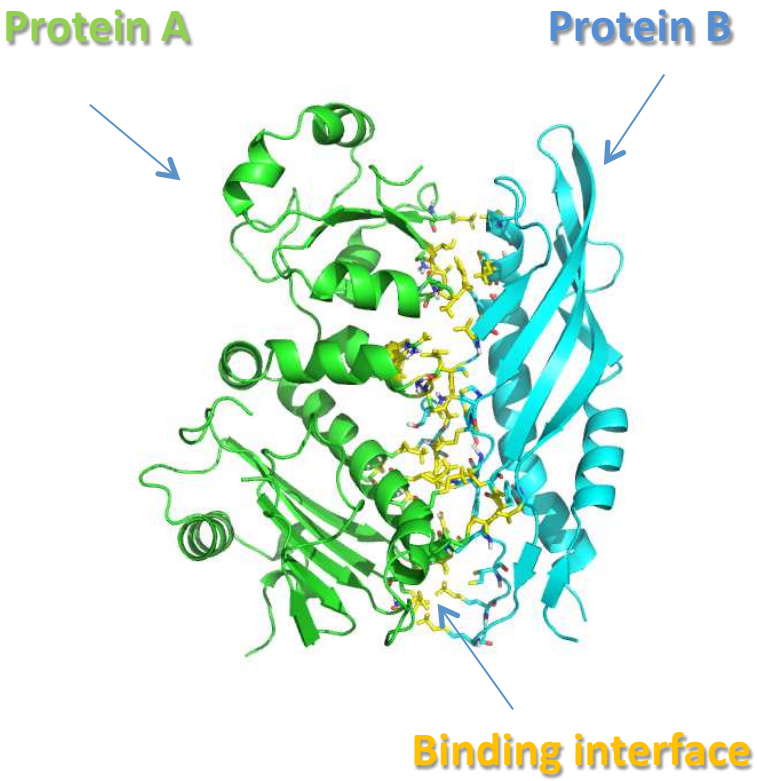
HADDOCK in the Grid



http://haddock.chem.uu.nl/enmr/haddock-tutorial.php

Demo

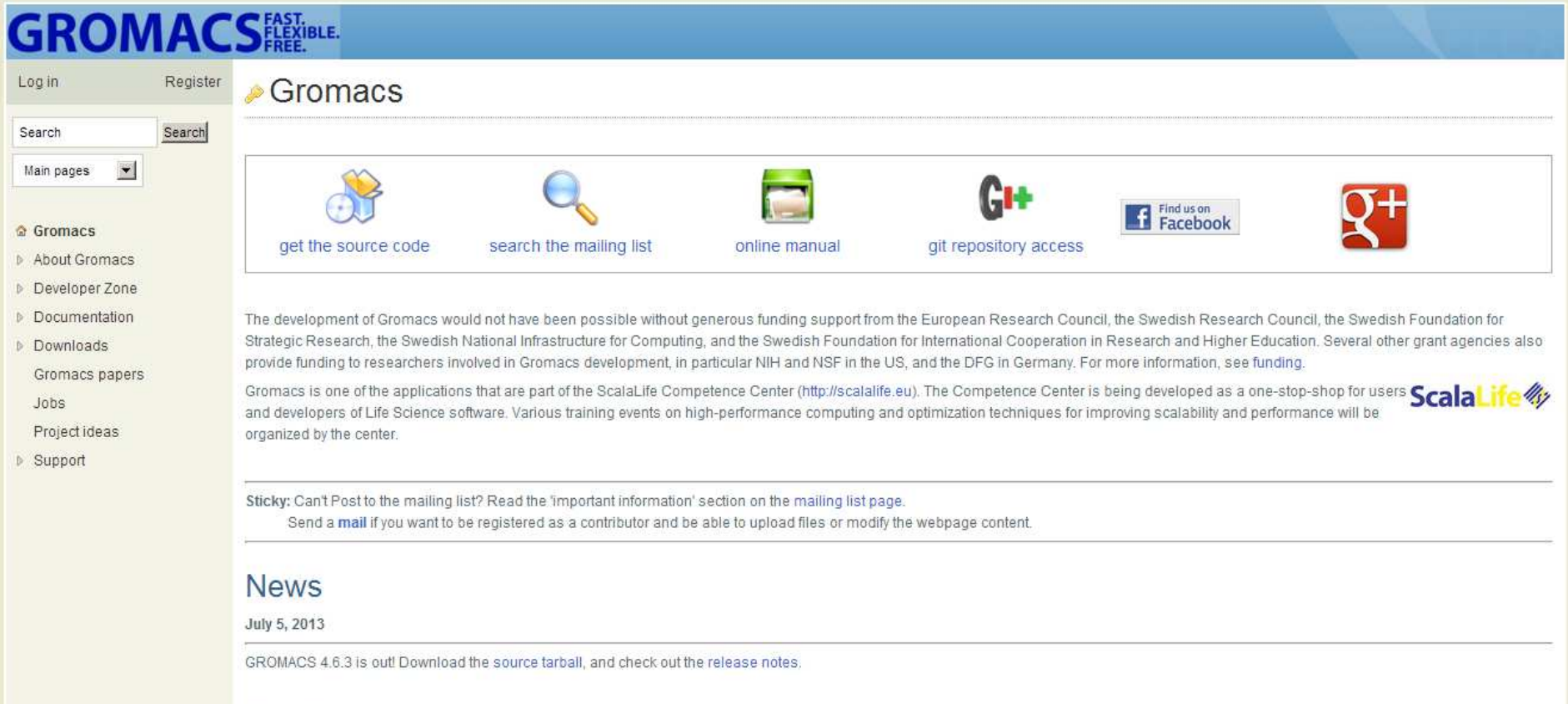
Protein – Protein interactions
HADDOCK in the Grid



Protein – Protein interactions

GROMACS

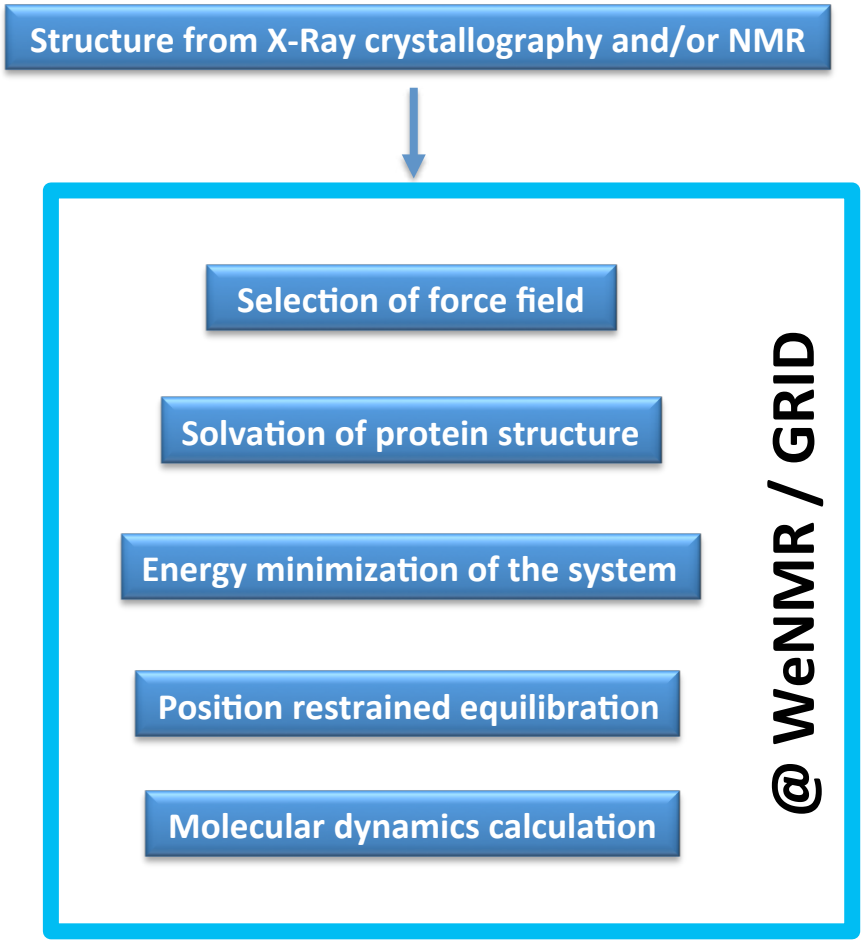
”perform molecular dynamics, i.e. simulate the Newtonian equations of motion for systems with hundreds to millions of particles.”



http://www.gromacs.org/

Protein dynamics

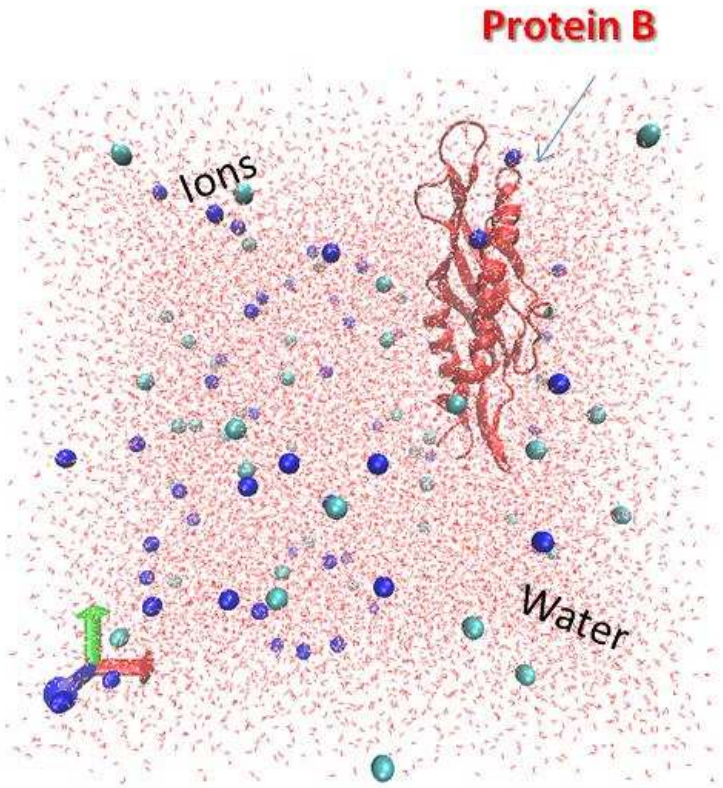
GROMACS in the Grid



http://www.wenmr.eu/wenmr/tutorials/nmr-tutorials/gromacs

Protein dynamics

GROMACS in the Grid



Structure from X-Ray crystallography and/or NMR

Selection of force field

Solvation of protein structure

Energy minimization of the system

Position restrained equilibration

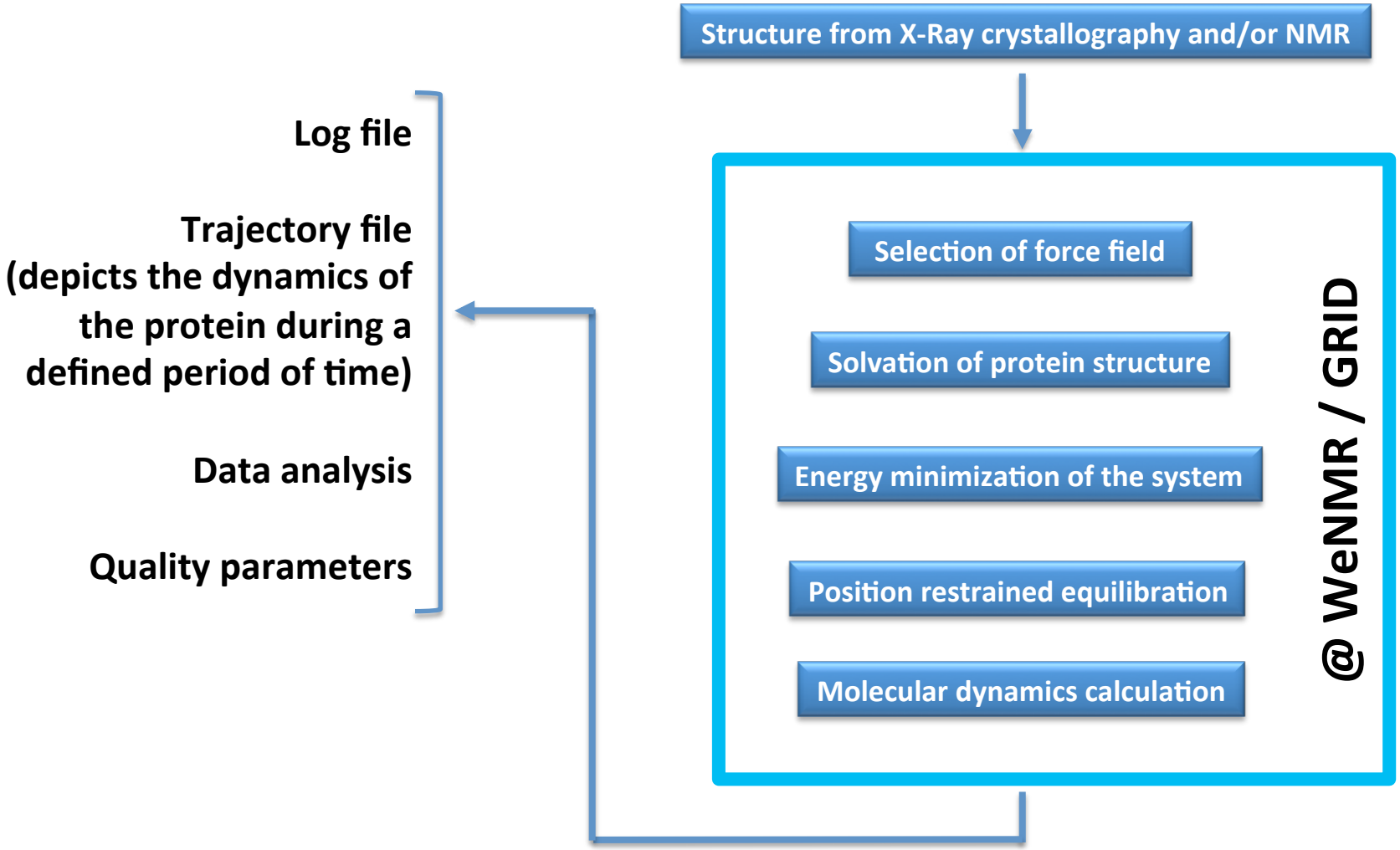
Molecular dynamics calculation

@ WeNMR / GRID

http://www.wenmr.eu/wenmr/tutorials/nmr-tutorials/gromacs

Protein dynamics

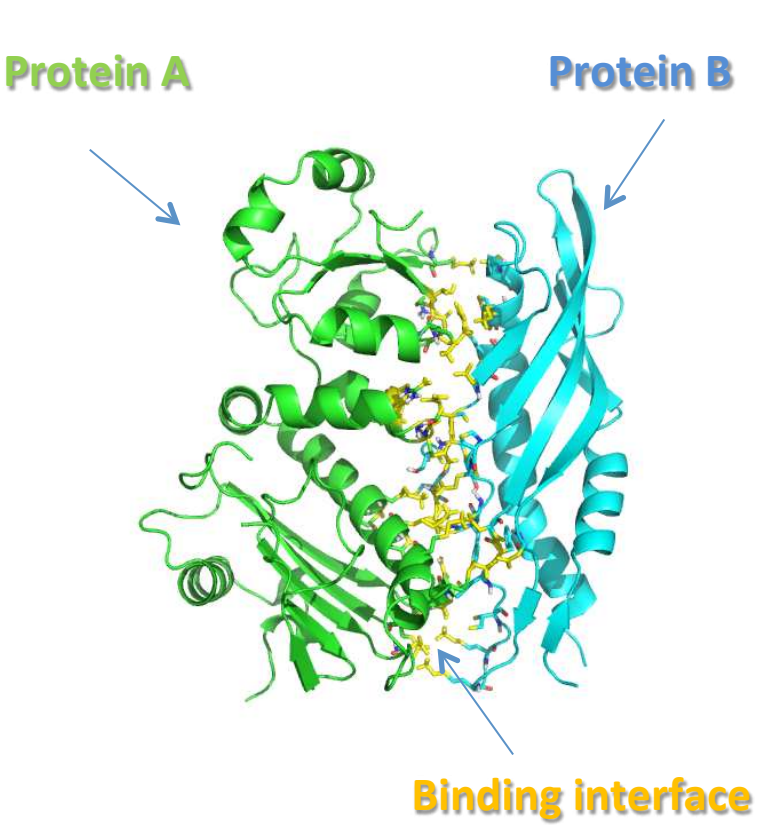
GROMACS in the Grid



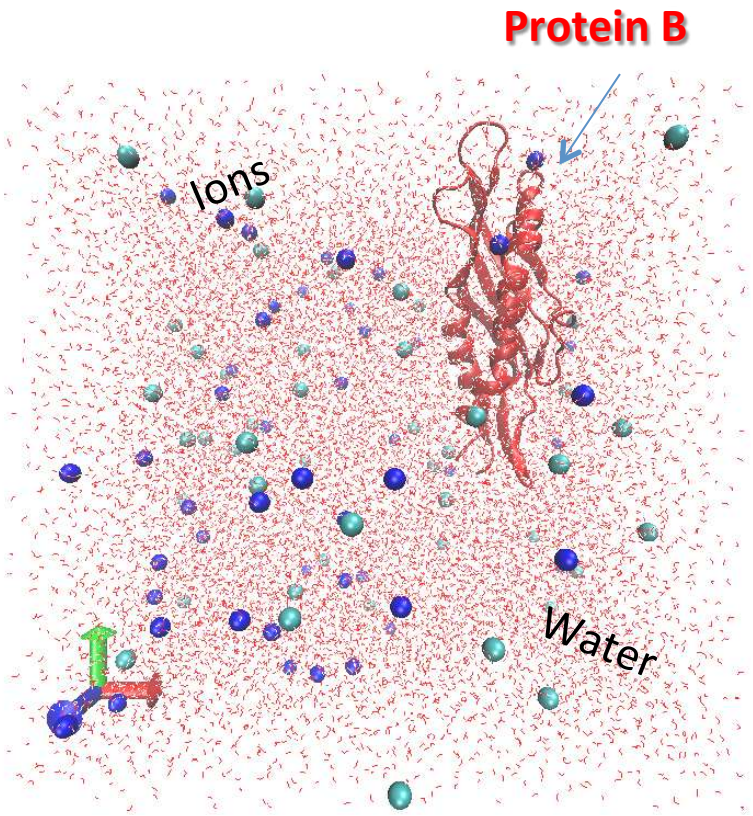
http://www.wenmr.eu/wenmr/tutorials/nmr-tutorials/gromacs

Demo

Two examples of the use of Grid application in Structural Biology

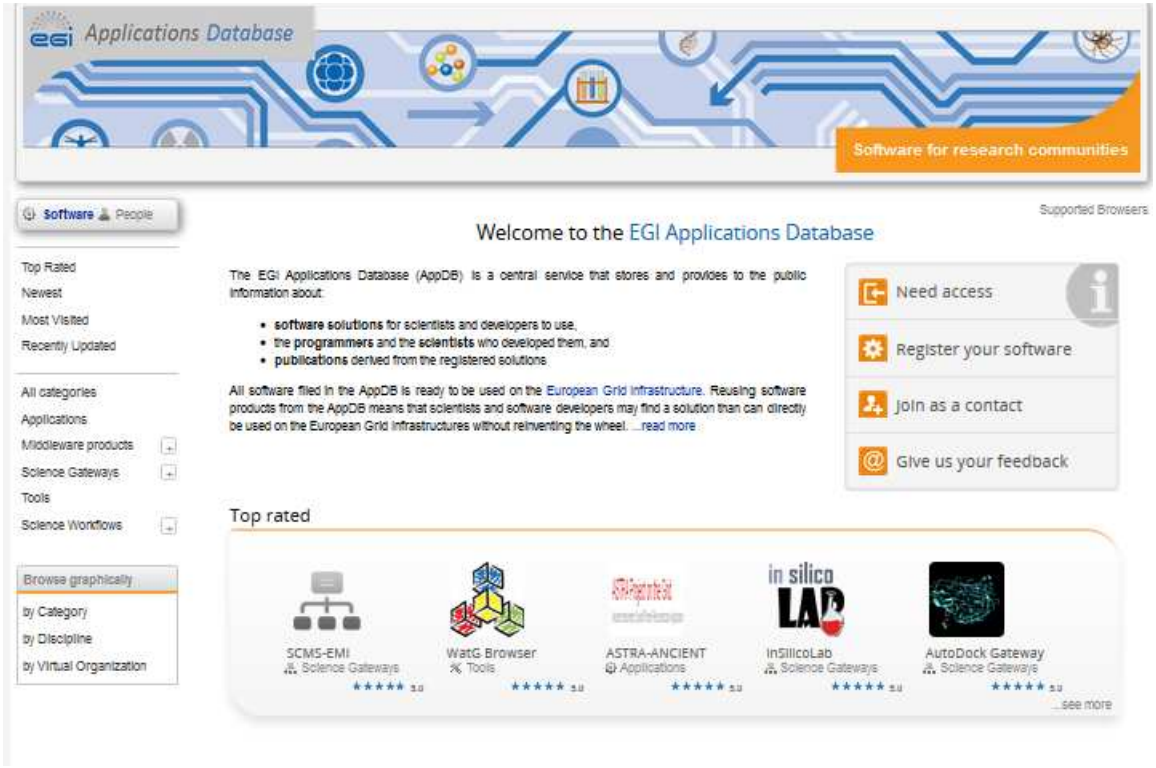


Protein – Protein interactions



Protein dynamics

More Examples/Application in Structural Biology and other fields



http://appdb.egi.eu/

https://www.egi.eu/case-studies/index.html

The EGI Champions

EGI Champions are enthusiastic scientists using grid computing for their research and keen to go to conferences and spread the word about the benefits of working with EGI.

Welcoming new users to the grid

Bridge users and developers

https://wiki.egi.eu/wiki/Champions
http://www.egi.eu/community/egi_champions/champions_network.html