Recherche de Nouveaux Médicaments sur grille

Jean Salzemann

Laboratoire de Physique Corpusculaire Clermont-Ferrand
IN2P3/CNRS

Problematic

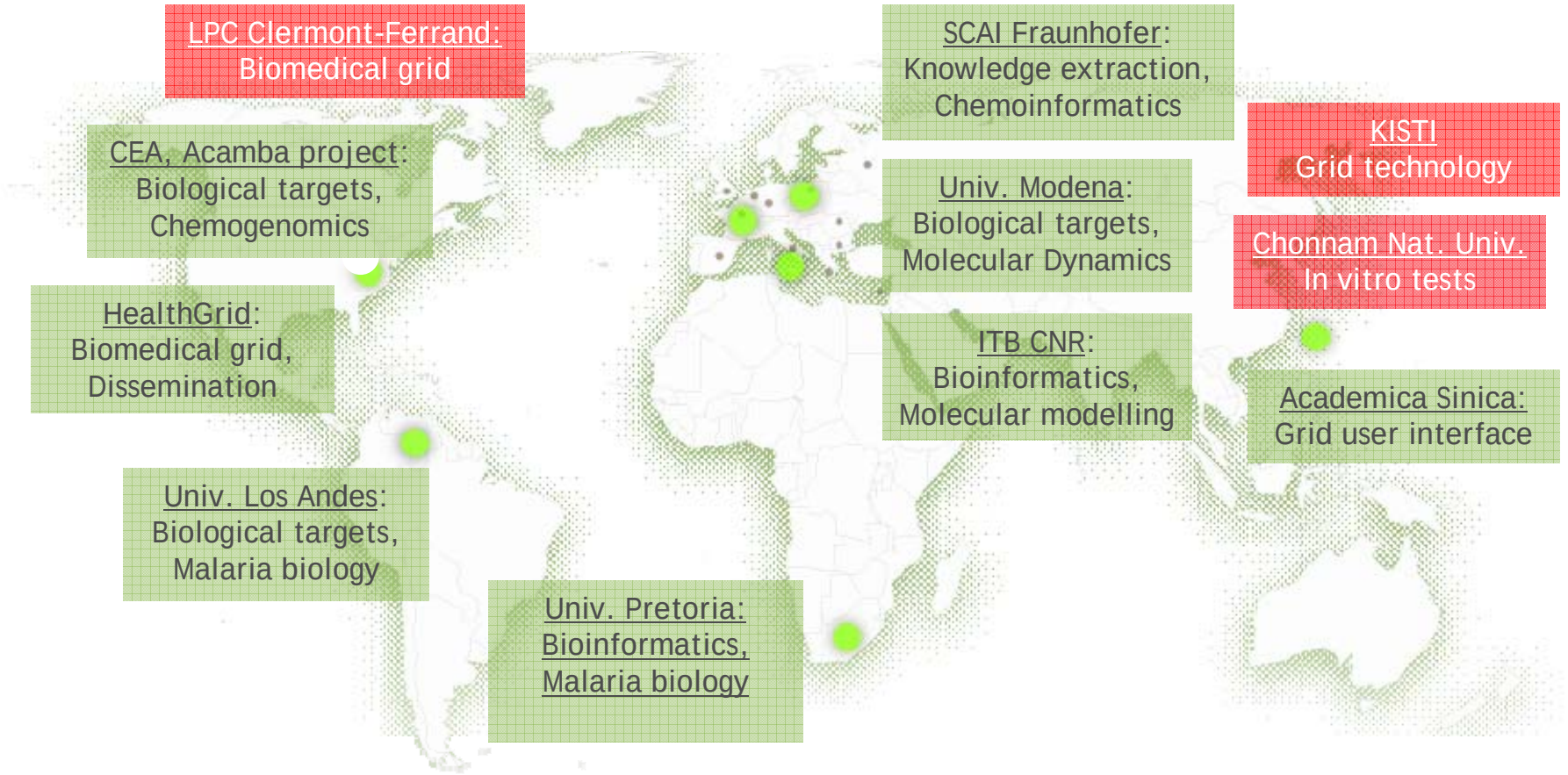
- Two statements
  - Algorithms related to life sciences are CPU intensive
  - In particular virtual screening can help drug development

- Added value of grid computing
  - The grid provides the **centuries of CPU** cycles required **on demand**

  - The grid provides the **reliable and secure** data management services to **store and replicate** the biochemical inputs and outputs

  - The grid offers a **collaborative environment** for the **sharing** of data in the research community on emerging and neglected diseases

WISDOM initiative



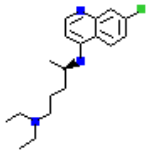
High-Throughput Virtual Screening

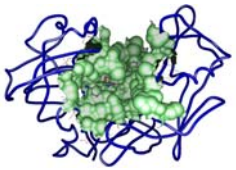
PCSV

Millions of chemical compounds available in laboratories

High Throughput Screening
1-10$/compound, nearly impossible

Chemical compounds (ZINC): 4.3 million

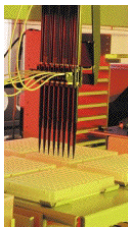




Molecular docking (FlexX)
~413 CPU years, 1.738 TB data

Data challenge on Grid
~90 days on ~5000 computers

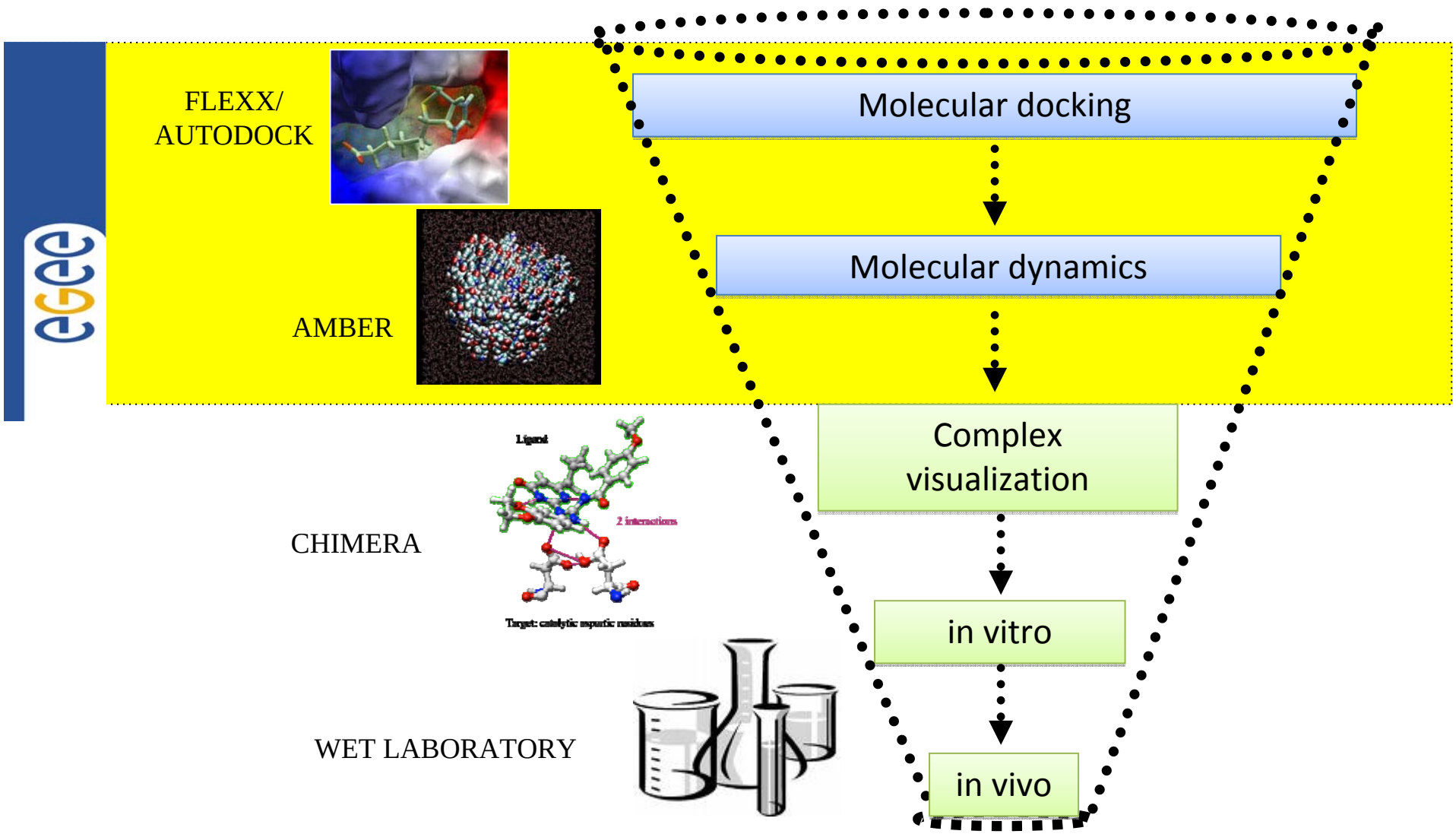
Targets (PDB): Plm, PvDHFR
PfDHFR, GST, tubulin



Lab testing

→ Leads
→ Clinical testing
→ Drug

Biological workflow



Grid results

- **To show the relevance of computational grids in biomedical applications**

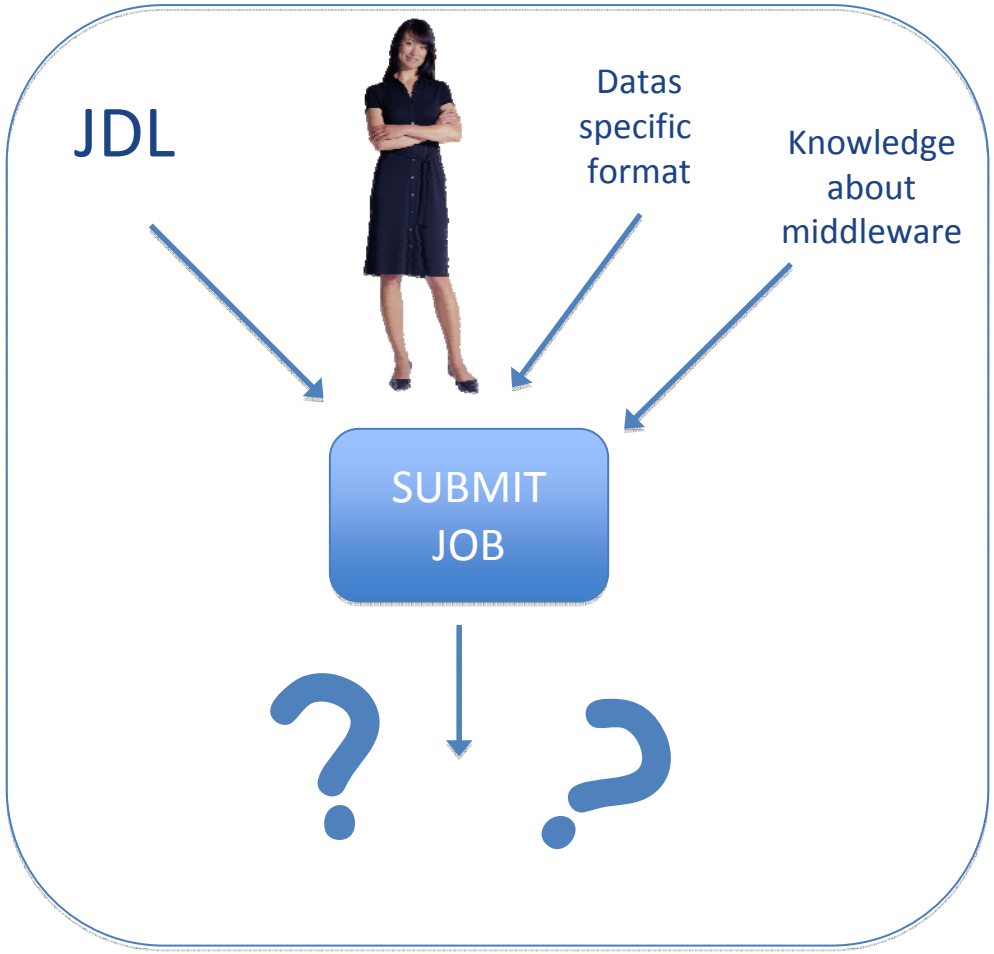
	Number of dockings	CPU years	Real Time	CPUs used	Produced Data size	Crunching Factor	Distribution efficiency
Malaria I	41 millions	80	6 sem	1700	1TB	400	25%
Malaria II	142 millions	400	2,5 mois	Jusqu'à 5000	1,6 TB	2000	40%
Avian Flu	4 millions	100	1,5 mois	1700	800 GB	900	50 % (>80% DIANE)
Diabetes	300, 000	40	2,5 jours	7000		6000	85 %

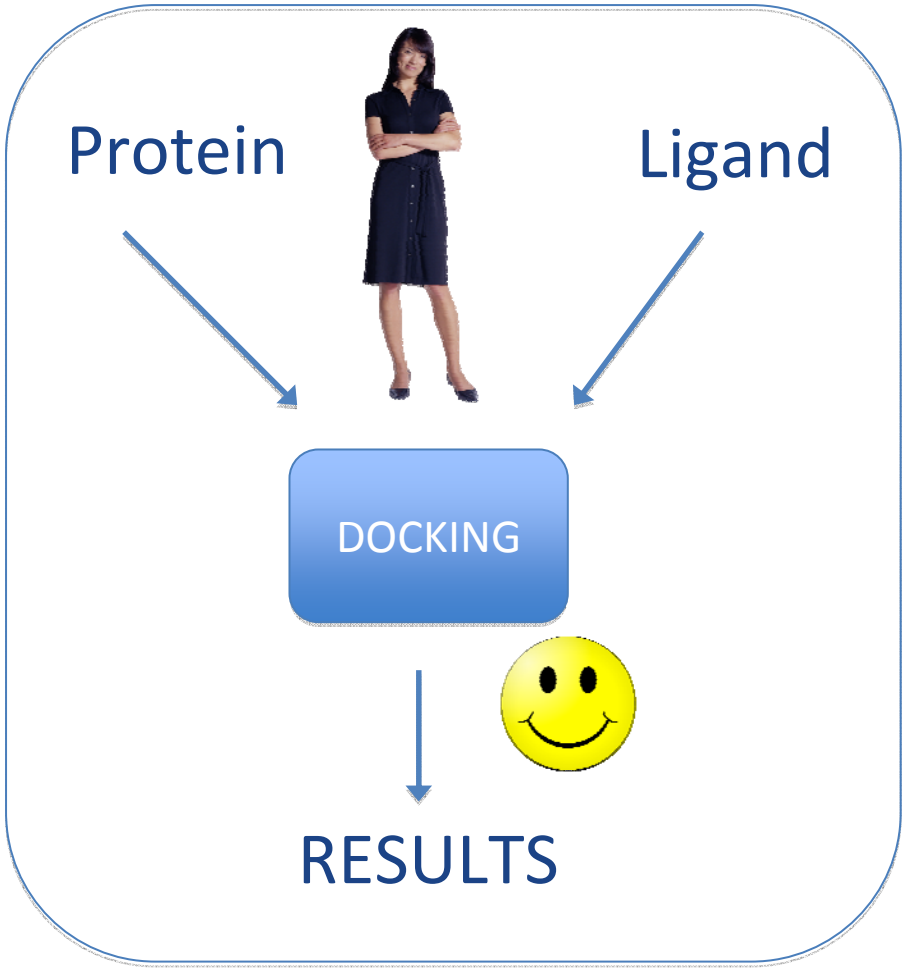
Wet Lab results

Prot.	Nbre composés testés	Type d'analyse	Référence	Nbre actifs	IC50	Ki
PM II	30	FRET	Pepstatin A IC50=4.3 nM	26 / 30	4.3 nM-1.8 µM	
Neuraminidase	185	Fluorogenic substrate	Oseltamivir	79/185 (59 > Oseltamivir)		
GST	32 26 (	Colorimétrie	S-hexyl glutathione Ki=35 µM	4 / 32		200-400µM

Conclusion

- The final goal of our activities is to hide the grid for the end user





- **To provide the grid to non-experts users**

WISDOM ENVIRONMENT

- Complex and unflexible
- For people familiar with GRID
- Drug discovery applications

2005-2008

TAKE ADVANTAGE OF THE GRID COMPUTING AND STORAGE RESOURCES

2008-…

- SOA
- Non-expert users
- Several bioinformatic's tools