Domain-oriented services and resources of Polish Infrastructure for Supporting Computational Science in the European Research Space – PLGrid Plus

InSilicoLab for chemistry

Joanna Kocot, Klemens Noga, Mariusz Sterzel

ACC Cyfronet AGH

Current Trends in Theoretical Chemistry VI , Kraków, 4 IX 2013

INNOVATIVE ECONOMY
NATIONAL COHESION STRATEGY

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND

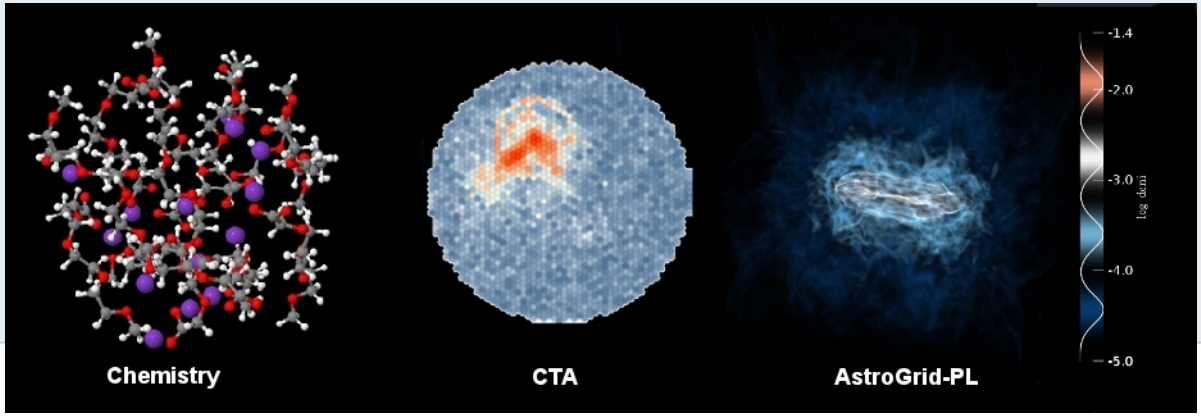
Outline

- InSilicoLab
  - features
  - technological background
- Grid accounts
  - grid certificates
  - Virtual Organisations
- Hands-on
  - grid certificate management
  - Quantum-Chemistry calculation
  - Trajectory Sculptor
  - experiment “workflows”

Overview

InSilicoLab - http://insilicolab.cyfronet.pl

- InSilicoLab - framework of application portals
  - enable execution of large-scale, long-lasting data- and computation-intensive experiments
  - support the management of complex calculations
  - automate repetitive cycles of calculations
  - enable joint analysis of results from massive parallel computations
- Scientific domains
  - quantum chemistry
  - Cherenkov Telescope Array (CTA)
  - multihydrodynamic for astrophysics



Overview

InSilicoLab - http://insilicolab.cyfronet.pl

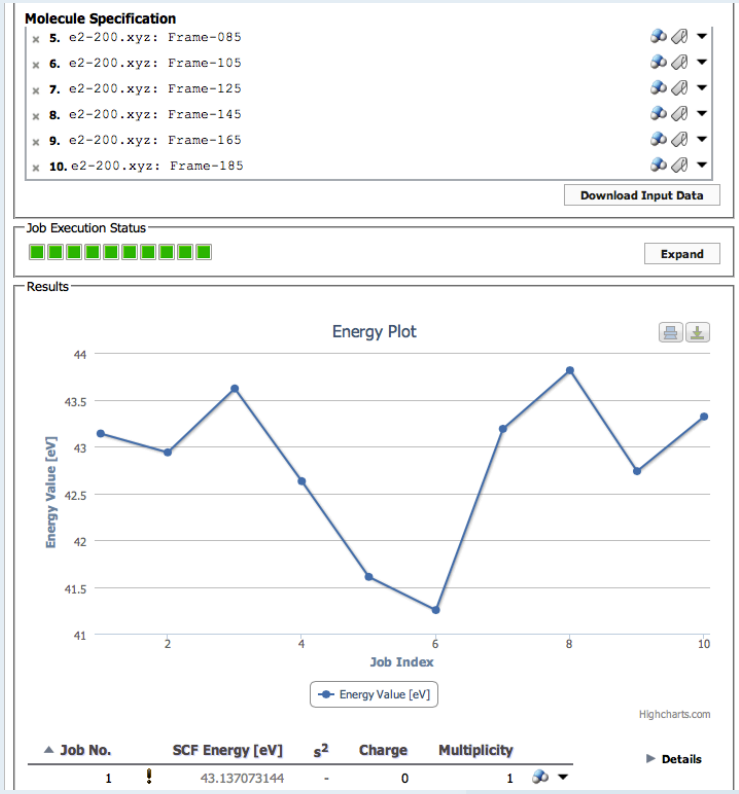
- Complete workspace in **web browser** for Chemist who wishes to perform Quantum Chemistry calculations
- Assist with
  - preparing input to various QC applications
  - performing computations on grid infrastructure
  - controlling complex and recurrent jobs
  - collecting output files
  - analysis of obtained results (also from many calculations at once)
- Main features
  - automatic parallelization
  - automatic results analysis
  - data categorization and visualization (JMol)
  - easy reuse of input and results in new experiments
  - automatic conformation scans (prototype)

Overview

InSilicoLab - http://insilicolab.cyfronet.pl

- Two types of experiments
  - Quantum chemistry calculations
  - Trajectory Sculptor

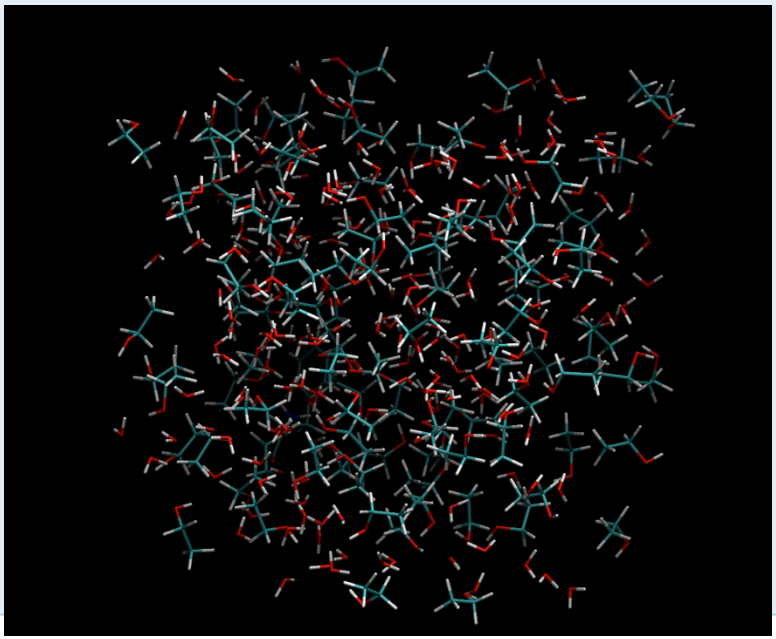
- **Quantum chemistry calculations**
  - QC applications – full integration
    - Gaussian
    - GAMESS
    - TUROBMOLE
  - automatic results analysis
  - simultaneous multiple geometry runs

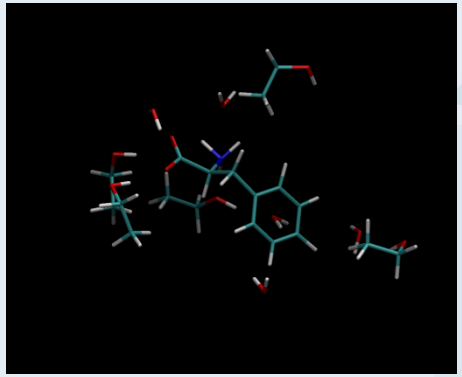


Overview

InSilicoLab - http://insilicolab.cyfronet.pl

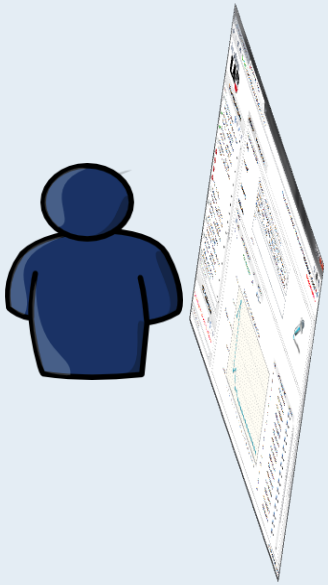
- **Trajectory Sculptor** - advanced tool for manipulation of Molecular Dynamics trajectories
  - automatic extraction of relevant parts of the input structure (based on user choices)
  - automatic processing frames chosen by user
  - results easily reused in quantum chemistry calculations





Technological background

InSilicoLab hides complexity of e-infrastructures



Experiment Management

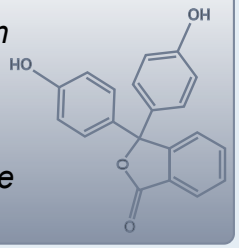
Specification
Execution
Browsing

Analysis
Sharing
Reuse
Classification

Result Management

Data Management

Preparation
Storing
Viewing
Provenance



Managed by the user

Experiment Logic

Automatic Parallelization

Data Structure

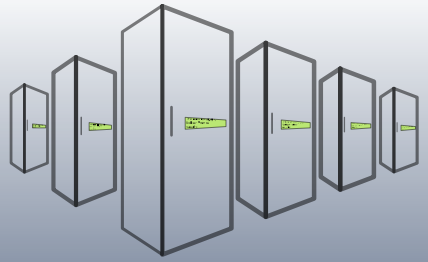
Provenance Tracking

Execution Engine

Metadata Model

Annotations Tagging

translation from domain- to resource-specific language



Computational Resources

Metadata Repository

Storage Resources

access to e-infrastructures

Technological background

Technology used

- Server
  - Java (jSAGA)
  - Ruby
- Middleware
  - gLite
  - DIRAC
- Computing Node
  - Quantum chemistry software packages
  - Python (cclib, lxml2)
- Web portal
  - Google Web Toolkit
  - jmol

Team

Joanna Kocot (leader)

Tomasz Szepeniec

Mariusz Sterzel

Maciej Golik

Tomasz Twaróg

Tomasz Grabarczyk

Piotr Wójcik

Michał Trzeciak

- Consultants (in chemistry): Andrzej Eilmes, Mariusz Sterzel, Klemens Noga

How to use InSilicoLab?

What is necessary to use InSilicoLab?

- Grid certificate (X.509 standard)
  - both PEM and PKCS#12 (.p12) formats supported
  - have to be installed in web browser
- Virtual Organisation membership
  - VOs available: gaussian, vo.plgrid.pl
- Some knowledge about Quantum Chemistry
  - familiarity with input formats of QC packages which user want to use (but we're working at Experiment Wizard and Quantum Chemistry Advisor)

Grid certificate

User identification and authentication

- Needed by all grid services to authenticate user and secure communication
- Protected with password (analogy: ATM card and PIN)
- Must be kept secure and not given to any other person
- Could be obtained from **certification authority** (**CA**)

- Certification authorities (CA) available in Poland
  - PL-Grid CA (https://plgrid-ca.pl)
  - Simple CA (http://plgrid-sca.wcss.wroc.pl)

- Most of grid services use short lived **proxy** certificates generated from user long term certificates due to security reasons
  - standard proxies are “fire & forget” type
  - **MyProxy** mechanism allows user renewable longer-time proxies

- Certificates could be used to sign and encrypt e-mails

Virtual Organisation

- Dynamic set of individuals or institutions defined around a set of resource-sharing on established conditions
  - research teams working on human genome, HEP (ATLAS, CMS, ALICE, etc.) projects
  - group of users using same computational software packages (e.g. Gaussian)
- Each member has to have grid certificate and sign up it to VO
- VOs available for chemists in InSilicoLab:
  - gaussian (sign up via https://voms.cyf-kr.edu.pl:8443/voms/gaussian/)
  - vo.plgrid.pl (sign up via https://portal.plgrid.pl/web/guest)

Hands on: certificate handling

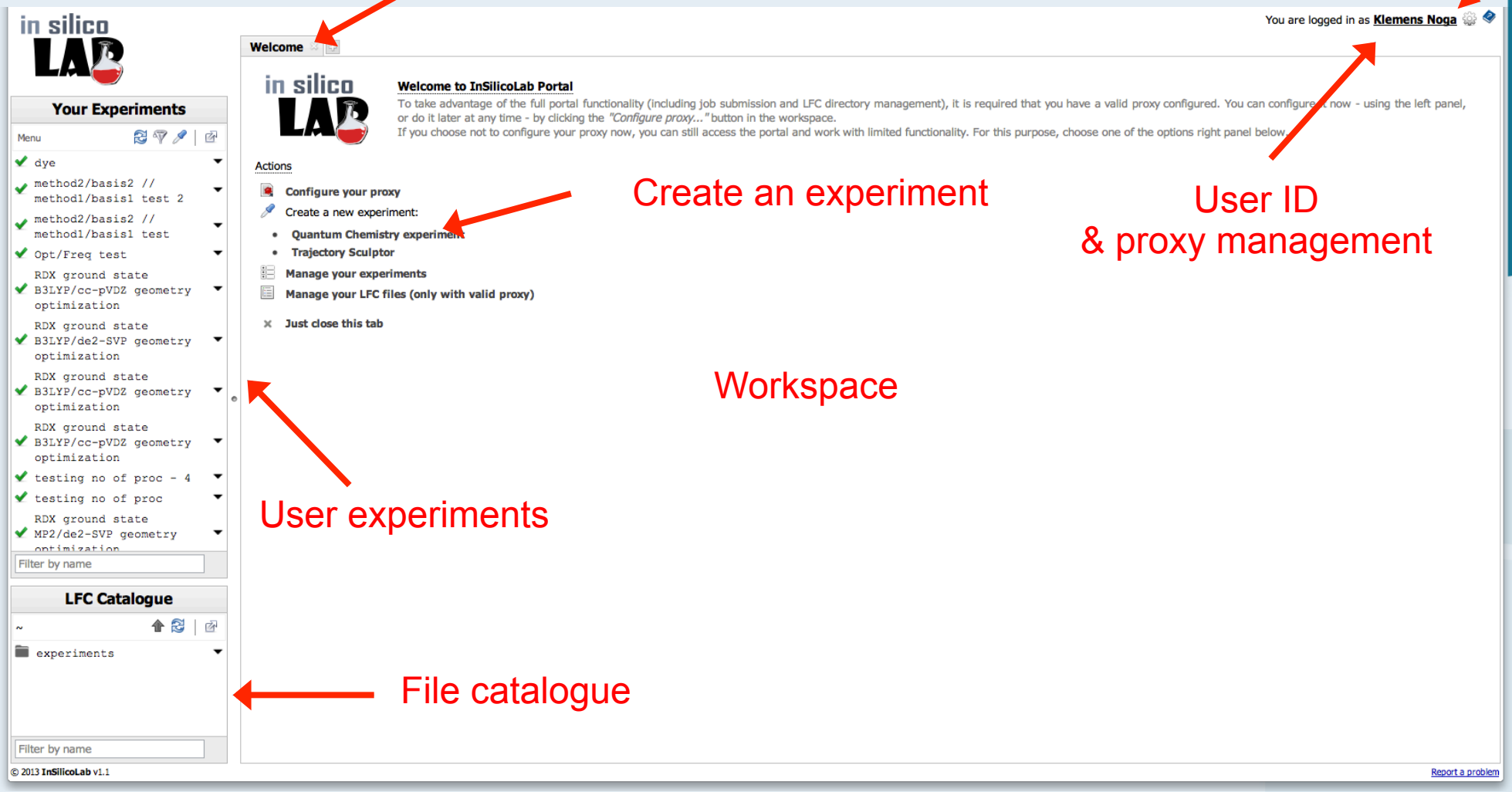
Grid certificate web browser import

- To access InSilicoLab web portal user needs to have valid grid certificate installed in web browser (or use PL-Grid OpenID)

- Import procedure (in Firefox)
  - use certificate in PKCS#12 (.p12) format
  - Linux: Edit → Preferences → Advanced → Security → Manage certificates → Your certificates → Import
    (pl: Edycja → Preferencje → Zaawansowane → Szyfrowanie → Wyświetl certyfikaty → Użytkownik → Importuj)
  - (Windows): Tools → Options → Advanced → Security → Manage certificates → Your certificates → Import
    (pl: Narzędzia → Opcje → Zaawansowane → Szyfrowanie → Wyświetl certyfikaty → Użytkownik → Importuj)
  - type in certificate password
  - done!
- Accounts on local computers: login: confXX, password: conf_XX; X∈{01,…,26}

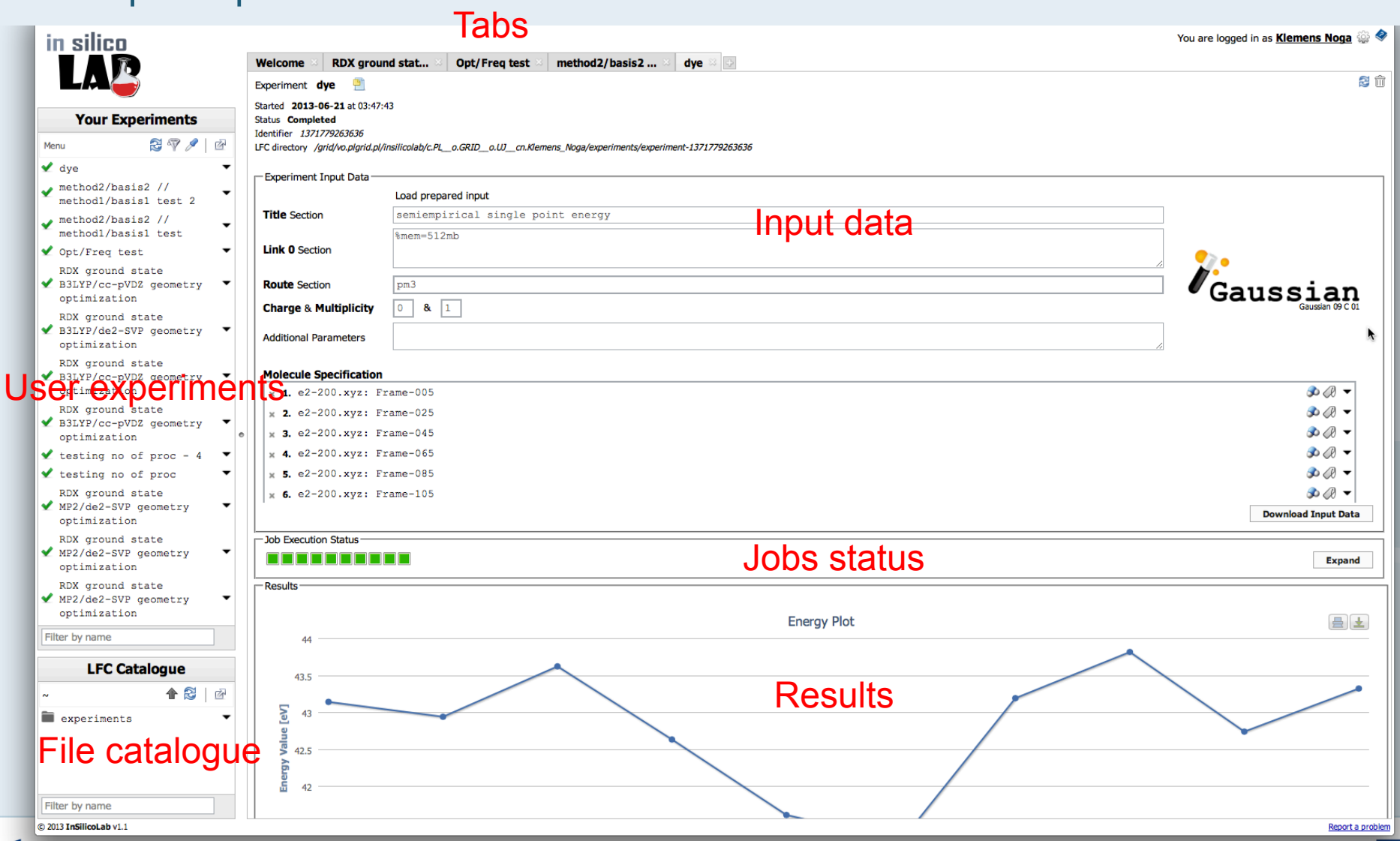
Hands on: portal structure

GUI



Hands on: portal structure

Example experiment



Hands on: Proxy management

Certificate upload and proxy creation

- Open **Configure your proxy**
  - upload your .p12 certificate
  - choose your favourite VO
  - check MyProxy box (for long renewable proxy usage)
  - type in passphrase
  - hit “Generate proxy” button

- You are ready for *in silico* experiments!

Hands on: Quantum Chemistry run

Simple QC Gaussian run

- Create new **Quantum Chemistry experiment**
  - upload rdx.pm3.gjf
  - view uploaded molecule geometry
  - upload another conformation (i.e. rdx.aae.xyz or rdx.all.xyz)
  - check data consistency
  - click “run” .-)

- After while
  - check which conformation has got lower energy
  - view optimal geometry
  - download detailed Gaussian log file
  - reuse lower energy structure to optimize geometry using B3LYP/cc-pVDZ

Hands on: Trajectory Sculptor

Trajectory Sculptor run

- Create new **Trajectory Sculptor experiment**
  - upload MD trajectory from md.2-200.xyz.b2z
    - Periodic Boundary Conditions with 44.7302 Å box
  - preview selected frame
  - find molecules in mixture
    - solute's atoms are in 1-38 range
    - there are 3 types of solvent molecules
      - Acetonitrile (atoms sequence CNCHHH)
      - perchlorate anions (atoms sequence ClOOOO)
      - lithium cations (Li)
    - preview findings
  - propose trimming conditions
    - you could use different metrics for different solvents
    - use closest molecule metrics due to obtain electro-neutral structure (i.e. 20 acetonitrile molecules, and one of each iones; syntax: 20;1;1)
    - preview findings

Hands on: Trajectory Sculptor

Trajectory Sculptor run cont.

- define frames range to be processed (there are 200 fames in MD file)
- reuse obtained trimmed frames in QC experiment (i.e. PM3 single point in Gaussian)
  - review obtained results
  - reuse selected frames using better QC method

Contact

http://insilicolab.cyfronet.pl

insilicolab@cyfronet.pl