



Project Number: **RI-312579**

Project Acronym: **ER-flow**

Project Full Title:

**Building an European Research Community through
Interoperable Workflows and Data**

Theme: **Research Infrastructures**

Call Identifier: **FP7-Infrastructures-2012-1**

Funding Scheme: **Coordination and Support Action**

Annex B – Computational Chemistry Community

Due date of deliverable: 28/08/2013

Start date of project: 01/09/2012

Actual submission date: --/--/2013

Duration: 24 months

Lead Contractor: University of Westminster

Dissemination Level: PU

Version: 1.0





1 Table of Contents

Contents

1	Table of Contents.....	2
2	List of Figures	3
3	List of Tables	4
4	Status and Change History	5
5	Glossary	6
5.1	Workflows and Meta-workflows.....	7
5.1.1	MD Solvation	7
5.1.2	Galaxy MD	8
5.1.3	AutoDock Vina: ASDD	9
5.1.4	Galaxy Docking.....	10
5.1.5	NWChem: Spectroscopic analysis	11
5.1.6	NWChem: Spectroscopic benchmarking.....	13
5.1.7	NWChem: Parameter sweep workflow.....	14
5.1.8	NWChem: Transition state analysis	15
5.1.9	NWChem: TD-UNI	17
5.1.10	NWChem: SpecWSUNI	18
5.1.11	NWChem: PopulationUNI	20
5.1.12	Galaxy QM.....	22
5.1.13	NWChem: QM-MD.....	23

2 List of Figures

Figure 1, NWChem for molecular dynamics	7
Figure 2, NWChem for molecular dynamics	8
Figure 3, Autodock Vina	9
Figure 4, Galaxy Docking	10
Figure 5, Spectroscopic Analysis	11
Figure 6, Spectroscopic Benchmarking	13
Figure 7, NWChem parameter sweep workflow for functional/basis set scans as chemical scheme	14
Figure 8, NWChem meta-workflow for transition state (TS) analysis	15
Figure 9, TDUNI. NWChem workflow for TD-DFT in UNICORE	17
Figure 10, NWChem meta-workflow for spectroscopic analysis in UNICORE	18
Figure 11, NWChem workflow for full population analysis using NWChem and NBO, AOMix and AIM in UNICORE	20
Figure 12, Quantum chemical workflow using Galaxy and NWChem	22
Figure 13, Protein molecular dynamics equilibration workflow	23
Figure 14, Meta-metaworkflow combining the two domains with spectroscopic metaworkflows	23

3 List of Tables

Table 1, Deliverable Status	5
Table 2, Deliverable Change History	5
Table 3, Glossary	6
Table 4, MdSolvation technical details	7
Table 5, GalaxyMD technical details	8
Table 6, ASDD technical details	9
Table 7, Galaxy docking technical details.....	10
Table 8, Spectroscopic Analysis technical details	12
Table 9, Spectroscopic Benchmarking technical details	13
Table 10, Transition state analysis technical details	16
Table 11, TD-UNI technical details.....	17
Table 12, SpecWSUNI technical details	19
Table 13, PopulationUNI technical details	21
Table 14, GalaxyQM technical details	22
Table 15, QM-MD technical details	23

4 Status and Change History

Status:	Name:	Date:	Signature:
Draft:	Gabriele Pierantoni	16/08/2014	n.n. electronically
Reviewed:			n.n. electronically
Approved:			n.n. electronically

Table 1, Deliverable Status

Version	Date	Pages	Author	Modification
0.1	29/09/14	All	GP	Created

Table 2, Deliverable Change History



5 Glossary

SSP	SHIWA Simulation Platform
WP	Work package

Table 3, Glossary

5.1 Workflows and Meta-workflows

5.1.1 MD Solvation

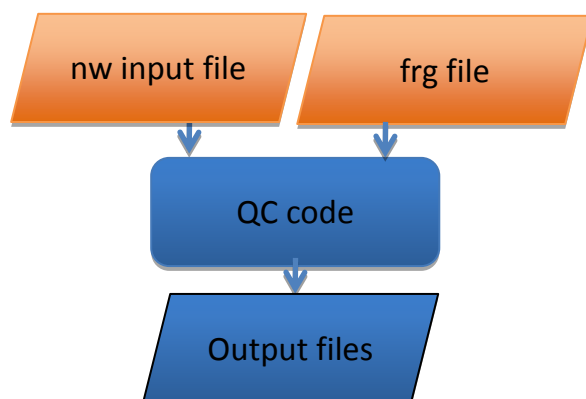


Figure 1, NWChem for molecular dynamics

5.1.1.1 Nature and Relevance

Molecules do mostly not exist in gas phase but in solution. The accurate modelling of the solvent environment is complicated and a fundamental challenge in computational chemistry. Multiple methods have been developed, e.g. the description of charge densities or probabilities of solvent molecules around the solute or of explicit molecules around the solute. These explicit molecules can be treated quantum mechanically which is computationally expensive and does not take into account the fluxional character of solvent-solute interactions. Moreover, these explicit molecules can be treated by a molecular dynamics approach where the solvent molecules are only treated molecular mechanically but they are allowed to move dynamically and find the best position. We target to provide to the MoSGrid users different types of solvent models. Since most users try to stay within one code, we have used the molecular dynamics module of NWChem. This code is widely used for several other workflows and hence, this workflow can be used in several further use cases (e.g. workflows from year 1).

5.1.1.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5748
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5748&filename=Application.MDsolvation.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download
Contact details	Dr. Jens Krüger, e-mail: krueger@informatik.uni-tuebingen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 4, MdSolvation technical details

5.1.2 Galaxy MD

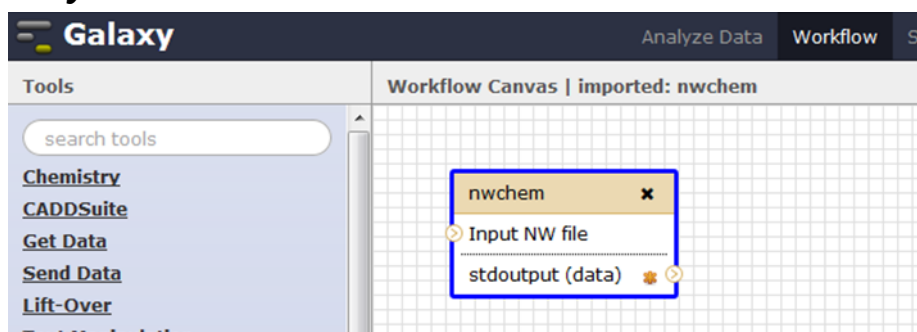


Figure 2, NWChem for molecular dynamics

5.1.2.1 Nature and Relevance

Molecules do mostly not exist in gas phase but in solution. The accurate modelling of the solvent environment is complicated and a fundamental challenge in computational chemistry. Multiple methods have been developed, e.g. the description of charge densities or probabilities of solvent molecules around the solute or of explicit molecules around the solute. These explicit molecules can be treated quantum mechanically which is computationally expensive and does not take into account the fluxional character of solvent-solute interactions. Moreover, these explicit molecules can be treated by a molecular dynamics approach where the solvent molecules are only treated molecular mechanically but they are allowed to move dynamically and find the best position. We target to provide to the MoSGrid users different types of solvent models. Since most users try to stay within one code, we have used the molecular dynamics module of NWChem. This code is widely used for several other workflows and hence, this workflow can be used in several further use cases (e.g. workflows from year 1).

The difference to the workflow “MD solvation” lies in the workflow engine. “MD solvation” is a pure WS-PRADE workflow whereas “Galaxy MD” is a Galaxy workflow. Galaxy is a workflow engine which is intensively used in the MD domain.

5.1.2.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5755
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5755&filename=Application.GalaxyMD.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download
Contact details	Dr. Jens Krüger, e-mail: krueger@informatik.uni-tuebingen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 5, GalaxyMD technical details

5.1.3 AutoDock Vina: ASDD

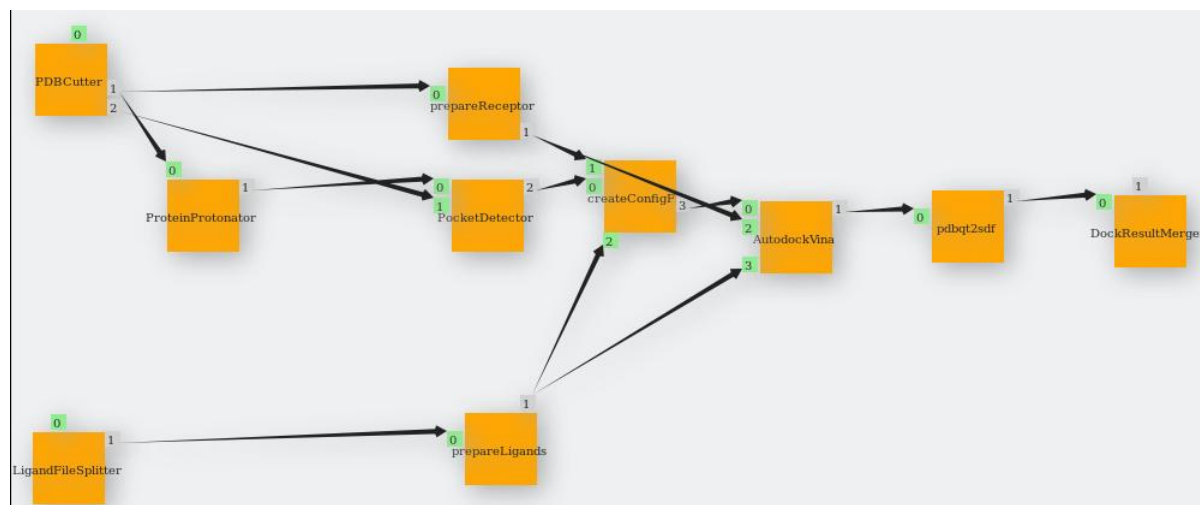


Figure 3, Autodock Vina

Contact information:

- Richard Grunzke, richard.grunzke@tu-dresden.de
- Dr. Jens Krüger, krueger@informatik.uni-tuebingen.de

5.1.3.1 Nature and Relevance

This workflow performs a docking procedure of ligands into a protein. It allows more advanced user input and is thus suitable for users who have performed docking before. The actual docking process includes splitting of the receptor-file into protein and reference ligand, docking with AutoDock Vina, and export of the dock results. AutoDock Vina creates a grid on the fly and thus no separate grid-building step is performed. For this workflow a protein structure file containing a protein in complex with a reference ligand and a file with the screening library (i.e., ligands that are to be docked into the protein) is required.

This workflow is based on the AutoDockVinaFull workflow. It additionally converts PDB to QT using Open Babel. QT is the AutoDock specific format that is able to much better handle the names of the ligands. In this case the workflow omits the parser to make the workflow better portable to other infrastructures.

5.1.3.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5747
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5747&filename=Application.AutoDockVinaASDD.pdf
User documentation	http://vina.scripps.edu/tutorial.html
Software documentation	http://vina.scripps.edu/
Contact details	Dr. Jens Krüger, e-mail: krueger@informatik.uni-tuebingen.de Richard Grunzke, e-mail: richard.grunzke@tu-dresden.de

Table 6, ASDD technical details

5.1.4 Galaxy Docking

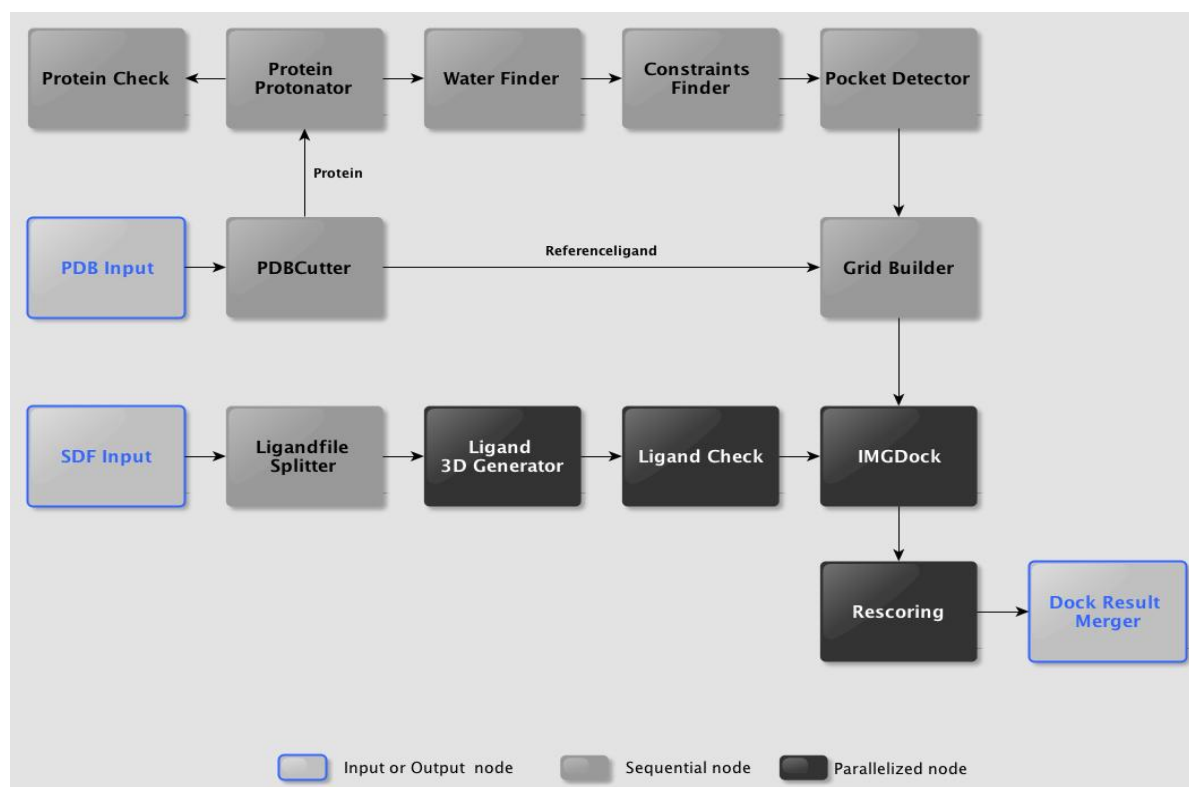


Figure 4, Galaxy Docking

5.1.4.1 Nature and Relevance

This workflow performs a docking procedure of ligands into a protein. This version is simplified so that only minimal user input is needed and thus is aimed at the novice user. The actual docking process includes splitting of the receptor-file into protein and reference ligand, grid building, docking and rescoring of the dock results. For this workflow a protein structure file containing a protein in complex with a reference ligand and a file with the screening library (ligands) is required.

5.1.4.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5743
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5743&filename=Application.GalaxyDocking.pdf
User documentation	http://www.ballview.org/Support/caddsuite-tutorial-1
Software documentation	http://www.ball-project.org/caddsuite
Contact details	Dr. Jens Krüger, e-mail: krueger@informatik.uni-tuebingen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 7, Galaxy docking technical details

5.1.5 NWChem: Spectroscopic analysis

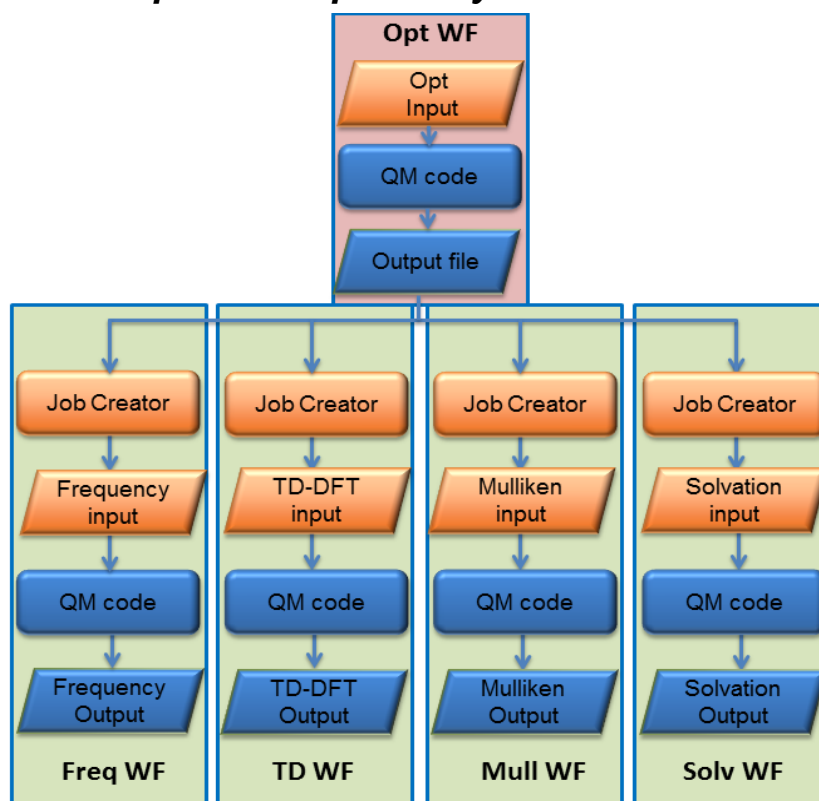


Figure 5, Spectroscopic Analysis

5.1.5.1 Nature and Relevance

As described in D5.2, we have developed a series of workflows using NWChem. They are highly useful for beginners who just want to perform a spectroscopic analysis. Spectroscopic analyses are widely used in organic and especially inorganic chemistry in order to simulate the UV/Vis absorption of a regarded molecule (TD job), its infra-red and/or Raman vibrations (frequency job), its relaxation in a solvent of choice (solvation job) and its population analysis (Mulliken job).

This analysis comprises a basic WF as optimisation followed by freq WF, TD WF, Mull WF and solv WF. A frequency workflow (freq WF) performs a frequency calculation, a step which is always needed to characterise a minimum as stable minimum. A time-dependent DFT workflow (TD WF) performs a TD-DFT calculation in order to obtain the optical properties of a molecule. The Mulliken workflow (Mull WF) calculates the Mulliken charges of every atom of a molecule. The solvent workflow (solv WF) arranges a solvent sheath around a molecule and performs a further optimisation step.

This workflow is so useful to quantum chemists that we have derived as Science Case of it.



5.1.5.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5739
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5739&filename=Application.SpectroscopicAnalysis.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download
Contact details	Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 8, Spectroscopic Analysis technical details

5.1.6 NWChem: Spectroscopic benchmarking

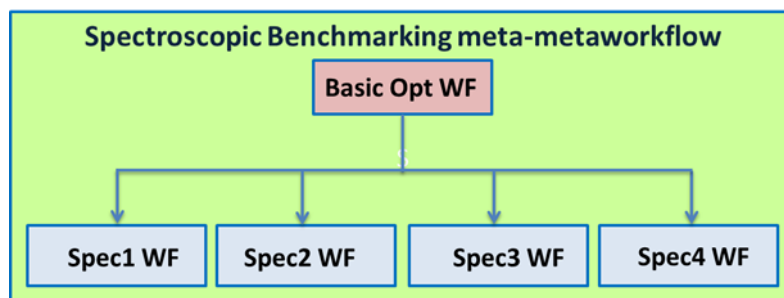


Figure 6, Spectroscopic Benchmarking

5.1.6.1 Nature and Relevance

The full simulation of molecular structures including the electronic structures comprises the calculation of optimized geometries, molecular orbitals, population analyses, frequencies, or optical absorptions. The combination of every one of these tasks as small basic workflow into a larger metaworkflow facilitates the simulators work enormously (see “Spectroscopic workflow”). With regard to real-life systems, the simulation has to tackle issues such as antiferromagnetic coupling between copper atoms, correct description of the coordination sphere and multiple conformations of the whole molecule. Methodologically, density functional theory is most appropriate here due to size of the system and investigated questions. Hence, the spectroscopic workflow needs to be performed several times for an array of functionals and basis sets which have to be tested for the ultimate structural and optical description with regard to experimental data.

Now, the spectroscopic metaworkflow can be combined into a new type of meta-metaworkflow with all being implemented in WS-PGRADE. Figure SpecBench shows that for example four spectroscopic workflows are combined after performing a basic optimization. This basic optimization (basic opt WF) serves as pre-optimization step which saves calculation time in all subsequent optimizations included in the spectroscopic workflows (specX WFs). A meta-metaworkflow saves a lot of time in this application - more than a normal meta-workflow.

5.1.6.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5745
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5745&filename=Application.SpectroscopicBenchmarking.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download
Contact details	Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 9, Spectroscopic Benchmarking technical details

5.1.7 NWChem: Parameter sweep workflow

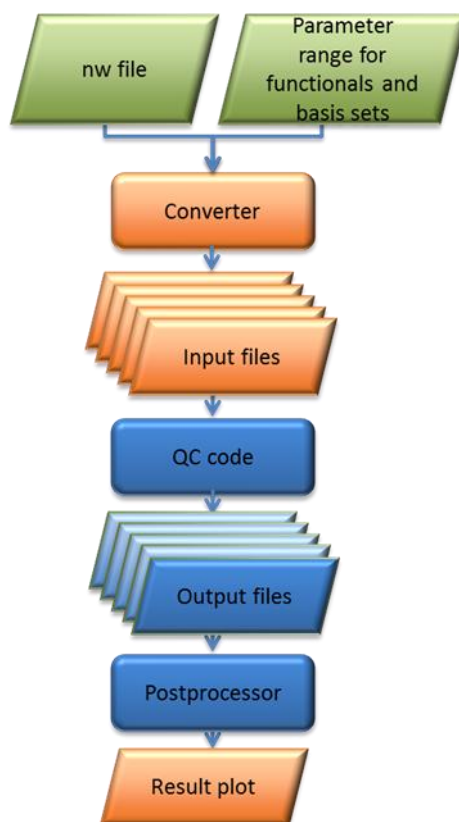


Figure 7, NWChem parameter sweep workflow for functional/basis set scans as chemical scheme

5.1.7.1 Nature and Relevance

In chemistry, it is useful that you test a large array of functionals and basis sets before you start with deeper orbital or spectroscopic analysis. Therefore, we have designed a parameter sweep workflow. Such a workflow would be highly useful for beginners and expert because the manual definition of input files is tedious. In fact, the submitted jobs are only optimisation jobs with different functional and basis sets.

The WS-PGRADE cookbook is very helpful but in this specific case, the chemical workflow developers did not manage to obtain meaningful output. The correct implementation of the scheme shown below requires an extremely high level of abstraction. Moreover the “parameters” are not numbers but different libraries of functionals and basis sets. Hence, the workflow was postponed. This workflow cannot be used at the moment since the implementation was too complicated for chemists.

5.1.8 NWChem: Transition state analysis

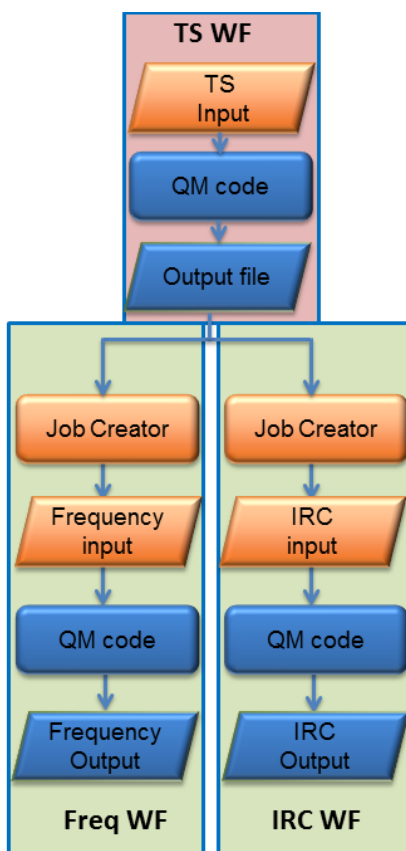


Figure 8, NWChem meta-workflow for transition state (TS) analysis

5.1.8.1 Nature and Relevance

As described in D5.2, we have developed a series of workflows using NWChem. In ..., we have described a spectroscopic meta-workflow. Here we have now a new combination of basic workflows with the transition state search workflow in D5.2.

This workflow is useful for more advanced users who want to perform a transition state search and subsequently want to analyse their transition state (TS). Herefore, one needs a frequency calculation (freq WF). This indicates by zero negative frequencies a normal minimum and by one negative frequency a saddle-point. The intrinsic reaction coordinate workflow (IRC WF) allows to search a reaction path downhill towards the educts and products of the TS.

As input file, only a TS.nw file is needed.

5.1.8.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5751
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5751&filename=Application.TransitionStateAnalysis.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download
Contact details	Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 10, Transition state analysis technical details

5.1.9 NWChem: TD-UNI

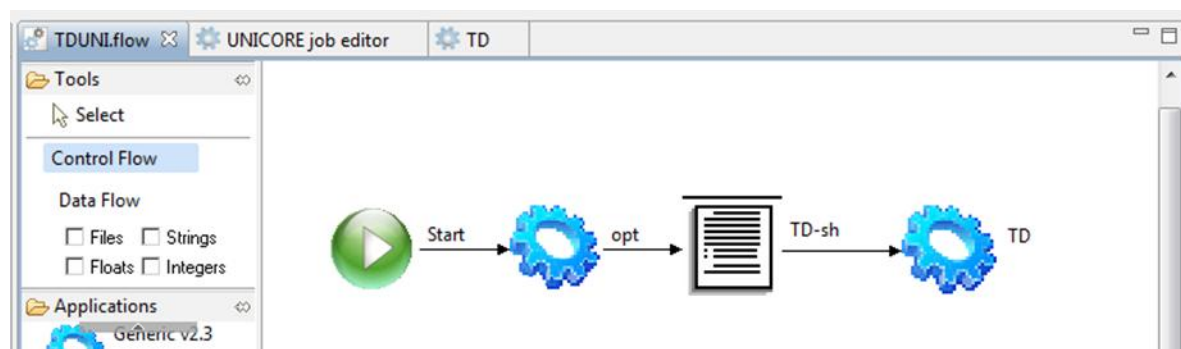


Figure 9, TDUNI. NWChem workflow for TD-DFT in UNICORE

5.1.9.1 Nature and Relevance

Time-dependent DFT serves for the calculation of optical transitions by ground state DFT. It replaces the complicated many-body time-dependent Schrödinger equation by a set of time-dependent single-particle equations whose orbitals yield the same time-dependent density. This is possible because the Runge-Gross theorem proves that, for a given initial wavefunction, particle statistics and interaction, a given time-dependent density can arise from at most one time-dependent external potential. This means that the time-dependent potential (and all other properties) is a functional of the time-dependent density. After the fundamental geometry optimisation, the TD-DFT step is performed in this workflow.

In this particular case, we have implemented it in UNICORE since this is directly running on several clusters without time delay. For larger calculations the UNICORE Rich Client offers good steering properties.

5.1.9.2 Further Technical Details

Information	URL
SHIWA Repository	Not yet.
Application description template	http://www.erflow.eu/documents/388575/775509/Application.TDUNI.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download .
Contact details	Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 11, TD-UNI technical details

5.1.10 NWChem: SpecWSUN/

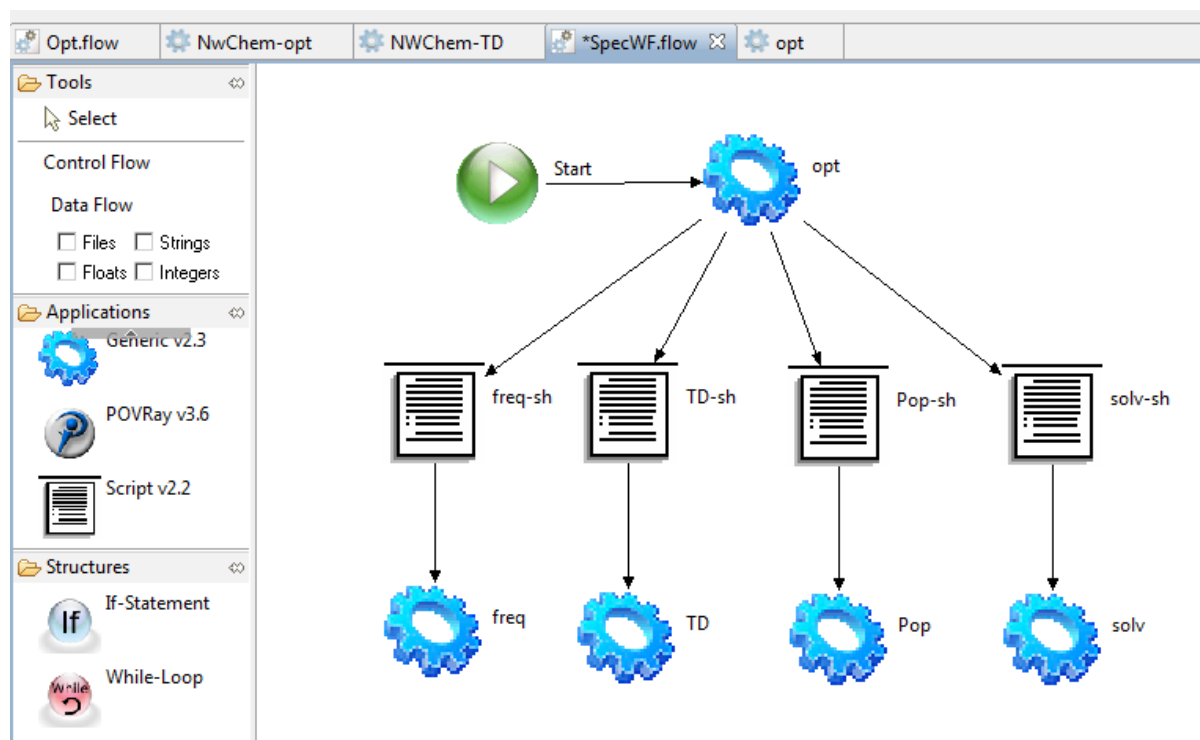


Figure 10, NWChem meta-workflow for spectroscopic analysis in UNICORE

Contact information:

- Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de
- Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

5.1.10.1 Nature and Relevance

As described in D5.2, we have developed a series of workflows using NWChem. They are highly useful for beginners who just want to perform a spectroscopic analysis. Spectroscopic analyses are widely used in organic and especially inorganic chemistry in order to simulate the UV/Vis absorption of a regarded molecule (TD job), its infra-red and/or Raman vibrations (frequency job), its relaxation in a solvent of choice (solvation job) and its population analysis (Mulliken job).

This analysis comprises a basic WF as optimisation followed by freq WF, TD WF, Mull WF and solv WF. A frequency workflow (freq WF) performs a frequency calculation, a step which is always needed to characterise a minimum as stable minimum. A time-dependent DFT workflow (TD WF) performs a TD-DFT calculation in order to obtain the optical properties of a molecule. The Mulliken workflow (Mull WF) calculates the Mulliken charges of every atom of a molecule. The solvent workflow (solv WF) arranges a solvent sheath around a molecule and performs a further optimisation step.

This workflow is so useful to quantum chemists that we have derived as Science Case of it. In this particular case, we have implemented it in UNICORE since this is directly running on several clusters without time delay. For larger calculations the UNICORE Rich Client offers good steering properties.



5.1.10.2 Further Technical Details

Information	URL
SHIWA Repository	Not yet.
Application description template	http://www.erflow.eu/documents/388575/775509/Application.SpecWSU.NI.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	. www.nwchem-sw.org/index.php/Download .
Contact details	Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 12, SpecWSUNI technical details

5.1.11 NWChem: PopulationUNI

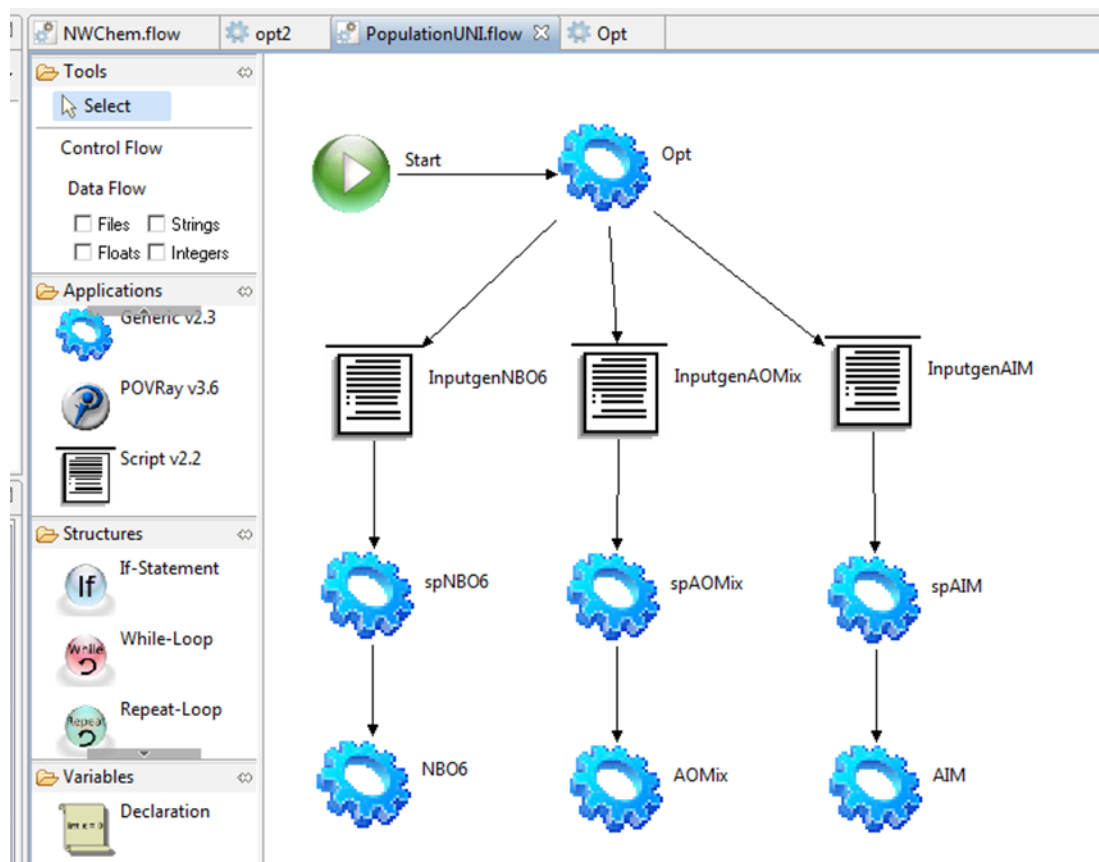


Figure 11, NWChem workflow for full population analysis using NWChem and NBO, AOMix and AIM in UNICORE

5.1.11.1 Nature and Relevance

Besides the standard natural bond orbital analysis implemented in many quantum chemical codes, the state of the art is represented by the NBO6.0 analysis. This is a standalone programme which needs special input. Hence, after the basic optimisation with NWChem a special input file for the subsequent single point job has to be generated by a script. This input file then goes into the single point calculation using NWChem (sp files). The output files of these calculations are then the input files for NBO6, AOMix and AIM. AOMix and AIM are small but commercial programmes where a working group licence is needed. They provide with different kinds of population analyses which allow for different orbital dissections, electronic analyses, calculation of bonding parameters and so on



5.1.11.2 Further Technical Details

Information	URL
SHIWA Repository	Not yet.
Application description template	http://www.erflow.eu/documents/388575/775509/Application.PopulationUNI.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	. www.nwchem-sw.org/index.php/Download .
Contact details	Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 13, PopulationUNI technical details

5.1.12 Galaxy QM

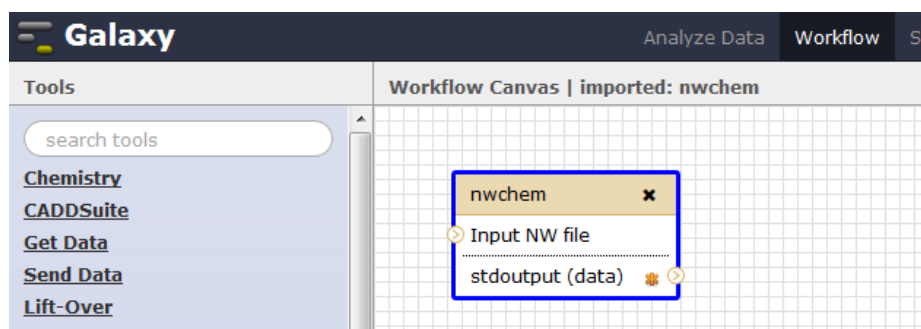


Figure 12, Quantum chemical workflow using Galaxy and NWChem

5.1.12.1 Nature and Relevance

This workflow performs a NWChem geometry optimisation of a desired molecule. It is highly useful for beginners who just want to perform a simple geometry optimisation but also as starting workflow for all upcoming type of jobs (Time-dependent DFT, solvent simulations etc.). The input file is a pre-prepared .nw file but for future applications we envision the direct usage of .xyz files as well. In order to work together with Galaxy users, we have transferred the basic optimisation workflow from D5.2 into Galaxy.

5.1.12.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5754
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5754&filename=Application.GalaxyQM.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download .
Contact details	Dr. Jens Krüger, e-mail: krueger@informatik.uni-tuebingen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de

Table 14, GalaxyQM technical details

5.1.13 NWChem: QM-MD

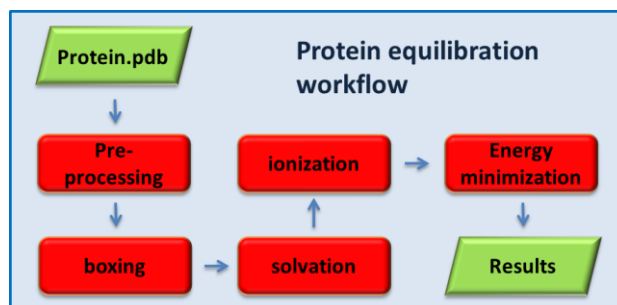


Figure 13, Protein molecular dynamics equilibration workflow

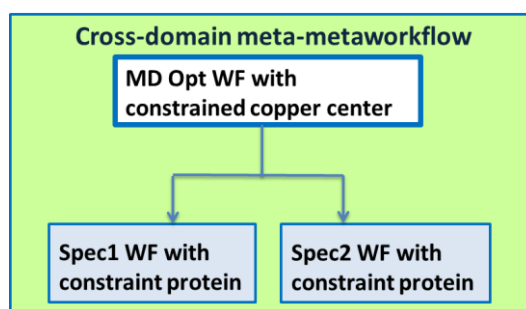


Figure 14, Meta-metaworkflow combining the two domains with spectroscopic metaworkflows

5.1.13.1 Nature and Relevance

If a true protein shall be modelled, mostly QM/MM procedures are applied. However, the correct description of the metal centres, their coupling and their reactivity remains problematic. Hence, a combination of molecular dynamics for the description of the protein environment is more appropriate. The inner centre containing the metal active site must still be modelled using quantum chemical methods such as density functional methods.

A typical molecular dynamics workflow is the protein equilibration (see also D5.2) which can be performed using Gromacs or NWChem. This MD workflow can be combined with the above described spectroscopic workflows to a cross-domain meta-metaworkflow depicted in Figure QM-MD .

5.1.13.2 Further Technical Details

Information	URL
SHIWA Repository	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/public/edit-application.xhtml?appid=5752
Application description template	http://shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/download?appid=5752&filename=Application.QM-MD.pdf
User documentation	http://www.nwchem-sw.org/index.php/Release62:NWChem_Documentation
Software documentation	www.nwchem-sw.org/index.php/Download .
Contact details	Dr. Alexander Hoffmann, e-mail: alexander.hoffmann@cup.uni-muenchen.de Prof. Dr. Sonja Herres-Pawlis, e-mail: sonja.herres-pawlis@cup.uni-muenchen.de Dr. Jens Krüger, e-mail: krueger@informatik.uni-tuebingen.de

Table 15, QM-MD technical details