West-Life Deliverable D2.1

Project Title:	World-wide E-infrastructure for structural biology
Project Acronym:	West-Life
Grant agreement no.:	**675858**
Deliverable title:	Interim report on T2.1 (Training programmes: user communities)
WP No.	2
Lead Beneficiary:	7 - Instruct
WP Title	Dissemination, Training and Outreach
Contractual delivery date:	30-April-2017
Actual delivery date:	30-April-2017
WP leader:	Instruct	7
Contributing partners:	STFC, NKI AVL, EMBL, MU, CSIC, CIRMMP, UU, INFN

Deliverable written by Narayanan Krishnan

Contents

1   Executive summary

Disseminating information and training of existing user communities is at the heart of the work package 2. Multiple user training events were conducted by West-Life partners enabling different communities to use the tools and services offered through the project. The events, conducted as workshops, focussed on management and analysis of structural biology data from different sources. In future, West-Life will continue supporting existing and new user communities in using specific services and their uses. Future training events will also enable dissemination of West-Life VRE. West-Life partners also contributed to user training events organised by collaborating projects/infrastructure. A list of these contributions is recorded and acknowledged in Appendix 6.

2   Project objectives

With this deliverable, the project has reached or the deliverable has contributed to the following objectives:

No.	Objective	Yes	No
1	**Provide analysis solutions for the different Structural Biology approaches**		N
2	**Provide automated pipelines to handle multi-technique datasets in an integrative manner**		N
3	**Provide integrated data management for single and multi-technique projects, based on existing e-infrastructure**		N
4	**Foster best practices, collaboration and training of end users**	Y	

3   Detailed report on the deliverable

3.1   Background

Training user communities, both new and existing, in tools and services offered through West-Life is key for the project to enable high impact research outcomes. The training events will be addressed towards the existing structural biology communities and new user communities, providing information and practical experience of data acquisition, handling and also provide information on availability and access to methods, tools and services. With these objectives in focus, West-Life has organised multiple training events, details of which are discussed in sections 3.2 and 3.3.

3.2   Training events (co-) organized by West-Life

A summary of training events and details are as below:

Title	Organiser	Country	Audience	Number of participants	Countries represented
Second Advanced methods for the integration of diverse structural data with NMR data	UU (Partner 8)	Netherlands	Structural Biology community	20	Czech Republic, France, Germany, Italy, Netherlands, Singapore, Spain UK, USA
First I2PC-FEI "hands on" course on image processing applied to the structural characterization of biological macromolecules	CNB-CSIC (Partner 5), FEI	Spain	Structural Biology community	20	Austria, Czech Republic, Denmark, France, Germany, Netherlands, Italy, Spain, UK
Cryo EM Solving the Structure of Macromolecul	Academia Sinica Grid Computi	Taipei	New user community	20	India, Malaysia, Netherlands, Spain,

ar Complexes: A Hands-on Workshop at International Symposium on Grids and Clouds 2016	ng Centre (ASGC), UU (Partner 8), CNB-CSIC (Partner 5)				Taiwan
From 2D images to 3D structures: A practical course on Electron Microscopy Single Particle Analysis	CNB-CSIC (Partner 5)	Spain	Structural Biology community	22	Czech Republic, Slovenia, Denmark, UK, France, Netherlands, Germany, Switzerland, Latvia, Italy, Spain
Cryo-EM Workshop at International Symposium on Grids and Clouds 2017	Academia Sinica Grid Computing Centre (ASGC), UU (Partner 8), CNB-CSIC (Partner 5)	Taipei	New user community	40	Czech Republic, France, India, Italy, Malaysia, Netherlands, Spain, Taiwan

3.2.1 Second Advanced methods for the integration of diverse structural data with NMR data

AIM: The course was organised by Utrecht University in collaboration with Instruct, WeNMR (http://www.wenmr.eu/), MoBrain (http://mobrain.egi.eu/) and BioExcel (http://www.bioexcel.eu/). The aims of this practical course were to train young researchers in the combined use of NMR data with other structural biology techniques, such as SAXS, X-ray diffraction or Cryo-EM in order to understand and rationalize the structural and dynamical properties of macromolecular systems. In addition, this course introduced and made extensive use of user-friendly e-science portals to facilitate the adoption of the methods and techniques by the students in their everyday work.

This course consisted of lectures describing the basics of the methods applied and hands-on training with the intuitive, web-based services of the WeNMR infrastructure, with the support of

SURFSara (https://www.surf.nl/en/about-surf/subsidiaries/surfsara/), which is part of the MoBrain CC and supported us in the organization of this course by providing both the location for the course as well as computational resources in the form of Cloud resources for the cryo-EM tutorial. The hands-on sessions allowed the students to get some experience with the various methods.

Twenty participants were selected from a total of 41 of applications, taking into account the following criteria: previous experience in the field of biomolecular NMR and structure biology; experience in computational aspects of structural biology; gender balance; relevance of current research to the course contents.

The participants were asked to complete an online survey at the end of the course. The overall evaluation was very positive. The separate topics of this course were assessed as well interrelated (60% gave the maximum grade of 5 for this), which highlights the integrative aspects of the course.

3.2.2 First I2PC-FEI "hands on" course on image processing applied to the structural characterization of biological macromolecules

AIM: The course was organised by the Instruct Image Processing Center at the CNB-CSIC (http://i2pc.cnb.csic.es), FEI Corporation (https://www.fei.com/) and Instruct, targeting both industrial and academic attendees, aiming to provide an in-depth presentation of current image processing methods in cryo-Electron Microscopy.

The course was focused on the structural analysis of purified macromolecules, a field normally referred to as "Single Particle Analysis". It was a very practical course and specially centered on the image processing steps leading to a cryo EM map; no prior knowledge on cryoEM image processing was assumed.

The I2PC personnel, together with some of the most qualified FEI personnel in cryoEM image processing, attended to the course providing a very close personal support. Attendees were encouraged to bring their own data (videos, images...).

The course was very well attended, with 20 attendees mostly from Instruct countries (17 out of 20) with a much-equalized gender balance (11 women, 9 men). Almost half the students came with their own data, although data quality was quite diverse among them. Still, 4 attendees ending up applying for Instruct Access projects to the I2PC, since data quality was good enough to deserve further analysis, while other students considered applying for EM access time through Instruct and iNEXT.

There were presentations both from FEI and from I2PC personnel, all of them very practical and covering aspects from EM sample preparation to image processing, including considerations about future developments in the field.

3.2.3 Cryo EM Solving the Structure of Macromolecular Complexes: A Hands-on Workshop at International Symposium on Grids and Clouds 2016

AIM: West-Life, through its partners Utrecht University and CSIC, conducted a hands-on workshop on CryoEM structure solution of macromolecular complexes at the International

Symposium on Grids and Clouds 2016. This workshop provided an introduction to image processing in 3DEM, mainly focused in single particle analysis (SPA). In this technique, we need to deal with several thousand projection images that are "views" of the same specimen under study. The final goal was to determine the spatial 3D orientations of these views and reconstruct a 3D model.

The International Symposium on Grids and Clouds (ISGC) 2016 (http://event.twgrid.org/isgc2016/index.html) was held at Academia Sinica in Taipei, Taiwan from 13-18 March 2016, with co-located events and workshops. The conference was hosted by the Academia Sinica Grid Computing Centre (ASGC).

The theme of ISGC 2016 focused on "Ubiquitous e-infrastructures and Applications". Contemporary research is impossible without a strong IT component – researchers rely on the existence of stable and widely available e-infrastructures and their higher-level functions and properties. As a result of these expectations, e-Infrastructures are becoming ubiquitous, providing an environment that supports large scale collaborations that deal with global challenges as well as smaller and temporal research communities focusing on particular scientific problems. To support those diversified communities and their needs, the e-Infrastructures themselves are becoming more layered and multifaceted, supporting larger groups of applications.

The goal of ISGC 2016 was to create a face-to-face venue where individual communities and national representatives can present and share their contributions to the global puzzle and contribute thus to the solution of global challenges.

3.2.4 From 2D images to 3D structures: A practical course on Electron Microscopy Single Particle Analysis

AIM: Electron Microscopy is experiencing a resolution revolution (Khulbrant, W., 2014, Elife 13:3), dramatically increasing its capacity to solve the structure of macromolecular complexes close to atomic resolution. The joint computational course between the Instruct Image Processing Center at the CNB-CSIC (http://i2pc.cnb.csic.es) and the Collaborative Computational Project for Electron cryo-Microscopy (CCP-EM, a member of CCISB) was established with focus on EM. The course ran for five days, from Monday lunch to Friday lunch, at the CNB-CSIC in Madrid. Tutors from Spain and the UK addressed the complete image processing workflow typical for macromolecular structural determination, from processing the raw images taken on the electron microscope to the interpretation of the derived 3D volumetric structure in terms of atomistic models.

The course made extensive use of the I2PC software integration framework Scipion, accessing the most widely used software suites in the field, such as EMAN, Relion, Frealign and XMIPP, as well as the CCP-EM/CCP4 program suite for the fitting of atomic structures.

The course was open to investigators at all levels. The participants worked with examples data, but they could bring their own data to look at if they wish.

The course was oversubscribed, so that we increased the target number of attendees by 10%, reaching 22. Most of the attendees were from Instruct countries, and about 1/3 of them from Spain. The gender balance was very good. Most attendees came from Academy, although 2 of them came from Industry, confirming the Industrial interest in cryoEM.

Everything worked as scheduled, and the presentations from I2PC and CCP-EM were highly complementary, with the I2PC focusing on obtaining the 3D map, and CCP-EM focusing on Modelling. We also had several presentations on the Electron Microscopy Data Base (EMDB) and related tools by EBI personnel.

Evaluations were positive, although the general feeling was that Program should have been extended over some more time.

3.2.5 Cryo-EM Workshop at International Symposium on Grids and Clouds 2017

AIM: West-Life partners Utrecht University and CSIC, conducted a Cryo-EM Workshop (http://event.twgrid.org/isgc2017/program/cryo-em-workshop) at the International Symposium on Grids and Clouds 2017. The workshop included practical and hands on training on tools and services used in Cryo-EM structure solutions (like Scipion, POWERFIT and DisVis) and scientific talks on current topics on Cryo-EM like Image processing problems and current perspectives and HR integrative modelling of complexes from fuzzy data.

The International Symposium on Grids and Clouds (ISGC) 2017 was held at Academia Sinica in Taipei, Taiwan from 5-10 March 2017, with co-located events and workshops. The main theme of ISGC 2017 is "Global Challenges: From Open Data to Open Science". The main purpose of open data is to improve sciences, accelerating specifically those that may benefit people. Nevertheless, to eliminate barriers to open data is itself a daunting task and the barriers to individuals, institutions and big collaborations are manifold. Open science is a step beyond open data, where the tools and understanding of scientific data must be made available to whoever is interested to participate in such scientific research. The promotion of open science may change the academic tradition practiced over the past few hundred years. This change of dynamics may contribute to the resolution of common challenges of human sustainability where the current pace of scientific progress is not sufficiently fast and where collaboration over e-Infrastructure will presumably help resolve the common problems of the people who are impacted. Access to global e-Infrastructure helps also the less globally organized, long-tail sciences, with their own collaboration challenges.

3.3   Training events to which West-Life partners contributed

West-Life training contributions are not just restricted to events organised by the project. Scientists from the project partners conduct sessions at workshops or courses organised by collaborators. These sessions are aimed at training users in specific methodologies/tools/services enabled through West-Life. In this report, we acknowledge 13 such tutorial sessions provided by West-Life partners (Appendix 6). These are workshops and tutorials (we do not count here presentations at general scientific conferences). Conducting training sessions at such events help the dissemination of West-Life project and tools/services enabled through the project to a worldwide community of users.

Appendix 1: Advanced methods for the integration of diverse structural data with NMR data

Course programme

April 11 (room 3.1): MS and solution NMR Morning: arrival
13:00-13:10 Welcome + intro (Bonvin)
13:10-13:30 Short intro to HPC/HTC infrastructure and services (SURFSara)
13:30-15:00 Heck: Characterization of large assemblies by mass spectrometry
15:00-15:30 coffee break
15:30-17:00 Boelens: Characterization of macromolecular assemblies by solution NMR

April 12 (room 3.3): NMR-based protein structure refinement + integration with SAXS
9:00-10:45: Parigi: Paramagnetic NMR data and integration of SAXS and NMR data for the analysis of protein structure/dynamics
10:45-11:15: coffee break
11:15:13:00: Rosato: Molecular dynamics for NMR-based protein structure refinement
13:00:13:45: Lunch
13:45-17:00: AMPS-NMR + FANTEN + MAXOCC computer practicals (Rosato / Parigi)

April 13 (room 3.3): Cryo-EM / Xray
9:00-10:45: Gros: Characterization of macromolecular complexes by X-Ray crystallography and cryo-EM
10:45-11:15: coffee break
11:15-13:00: Carazo: Cryo-EM analysis of biomolecular assemblies
13:00:13:45: Lunch
13:45:17:00: Cyro-EM computer practical (Jose Miguel de la Rosa)

Evening: workshop dinner

April 14 (room 4.1): Integrative modelling
9:00-10:45: Bonvin: Integrative modelling of biomolecular complexes by docking
10:45-11:15: coffee break
11:15-13:00: Bonvin: Integrative modelling of biomolecular complexes by docking
13:00:13:45: Lunch
13:45:17:00: HADDOCK computer practical (Bonvin + local assistant)

April 15 (room 3.3): Solid state NMR lecture + open discussion / question and answers
9:00-10:45: Baldus: Characterization of macromolecular assemblies by solid-state NM
10:45-11:15: coffee break
11:15-12:15: Open discussion / question and answers
12:15:13:00: Lunch Departure

Appendix 2: I2PC-FEI "hands on" course on image processing applied to the structural characterization of biological macromolecules

Course Programme

Monday, October 17th, 2016
12:00 - 12:30. Registration
12:30 - 13:30. Welcome Reception
13:30 - 13:45. Introduction to the Course (Jose Maria Carazo)
13:45 - 15:00. Cryo EM and 3D reconstruction (Jose Maria Carazo)
15:00 - 16:00. Basic Background: Fourier transform, central section theorem, sampling. (Carlos Oscar Sorzano)
16:00 - 16:30. Coffee break
16:30 - 1800. The Electron Microscope (EM): Direct Electron Detectors, image formation, aberrations, the Contrast Transfer Function (CTF) (Jose Maria Carazo)

Tuesday, October 18th
09:00 - 10:30. Single Particle workflow (Jose Miguel de la Rosa)
10:30 - 11:00: Coffee break
11:00 - 13:30. Cryo-TEM: From Evolution to Revolution – the Dawn of a New Scientific Age (past, present and future....) (Marc Storms)
13:30 - 14:30. Lunch
14:30 - 16:00. The single particle cryoTEM workflow - Practical aspects from sample preparation to automated image acquisition (Sonja Welsch)
16:00 - 16:30. Coffee break
16:30 - 18:00. Practical work: Video frame alignment, Contrast Transfer Function estimation, micrograph screening, particle picking and extraction (Jose Luis Vilas/Jose Maria Carazo)

Wednesday, October 19th
09:00 - 10:30. Practical work: 2D Classification Exploration (Pablo Conesa/Carlos Oscar Sorzano)
10:30 - 11:00: Coffee break
11:00 - 13:30. Practical work: Initial Volume (computational procedures and introduction to Randon Conical Tilt) (Pablo Conesa/ Carlos Oscar Sorzano)
13:30 - 14:30. Lunch
14:30 - 16:00. Practical work: 3D alignment and classification (Josue Gomez / Roberto Melero)
16:00 - 16:30: Coffee break
16:30 - 18:00. Practical work: 3D alignment and Reconstruction (Josue Gomez / Roberto Melero)
18: 00: Alcala de Henares city tour and Cocktail at the Parador.

Thursday, October 20th
09:00 - 10:30. Practical work: 3D validation and quality measures (Javier Vargas) 10:30 - 11:00. Coffee break
11:00 – 13:30. Introduction to model building (Roberto Marabini)
13:30 – 14:30: Lunch

Appendix 3: Cryo EM Solving the Structure of Macromolecular Complexes: A Hands-on Workshop

Monday, March 14th, 2016

Course Programme

09:00 - 10:30
- General introduction to Cryo-EM single particles technique.
- Introduction to Scipion framework

10:30 - 11:00
Coffee break

11:00 - 12:30
- Movies alignment (Xmipp)
- CTF estimation and screening (Ctffind4)
- Particle picking (Xmipp and Eman)

12:30 - 14:00
Lunch

14:00 - 15:30
- 2D classification (Spider)
- More picking with templates (Relion)

15:30 - 16:00
Coffee break

16:00 - 17:30
- 2D classification (Relion and Xmipp, maybe show precomputed results)
- Initial volume methods (Several packages)
- 3D classification and refinement (Relion, maybe show precomputed results)

Appendix 4: From 2D images to 3D structures: A practical course on Electron Microscopy Single Particle Analysis

Course Programme

<u>Tuesday, May 17</u>
08:30 – 09:00: Registration
09:00 - 09:15: Jose Maria Carazo: Introduction to the course
09:15 - 10:00: Jose Maria Carazo: Introduction to 3D EM
10:00 - 11:00: Carlos Oscar Sorzano: "Basic Background: Fourier transform, central section theorem, sampling"
11:00 – 11:30: Coffee break
11:30 - 12:00: Practical work. Jose Miguel de la Rosa: "Introduction to Single Particle workflow in Scipion".
12:00 - 13:00: Josue Gomez: "DDD, Screening micrographs and CTF"
14:00 - 15:00: Practical work. Josue Gomez: "particle picking and screening"
15:00 - 16:00: Practical work. Jose Miguel de la Rosa: "2D classification"
16:30 - 17:30: Practical work. Javier Vargas: "Initial volume"
17:30 - 18:30: Practical work. Roberto Melero: "Alignment, classification, and reconstruction - I"
<u>Wednesday, May 18</u>
09:00 - 10:30: Practical work. Roberto Melero: "Alignment, classification, and reconstruction - II"
11:00 - 12:00: Practical work. Carlos Oscar Sorzano: "Validation and analysis"
12:00 - 13:00: Presentation. Pablo Conesa: "Scipion Web Tools"
14:00 - 15:00: Lecture: Ardan Patwardhan: "EMDB and PDB"
15:00 - 16:00: Martyn Winn: "Overview of techniques for fitting and refining atomic models"
16:30 - 18:30: Agnel Praveen Joseph: "Scoring and assessment of model fitting"
18:30 - 19:30: Practical work. Agnel Praveen Joseph: "Rigid body fitting and scoring of multi-component assemblies"
<u>Thursday, May 19</u>
09:00 - 10:00: Tom Burnley: "Introduction to CCP-EM software "
10:00 - 11:00: Martin Wynn: "Refinement of atomic models against EM density using Refmac"
11:30 - 13:00: Practical work. Agnel Praveen Joseph: "Flexible fitting with FlexEM" 13:00 –
14:00 - 14:30: Tom Burnley: "High resolution: model building in Coot"
14:30 - 16:00: [Practical] Martyn/Tom: "Model building in Coot"
16:30 - 18:00: Practical work. Tom Burnley: "Model refinement with Refmac"
18:00 - 18:30: Joan Segura: "Incorporating Bioinformatics Annotations: 3DBionotes"
<u>Friday, May 20</u>
9:00 - 10:00: Ardan Patwardhan: "Volume matching with SMaSB / PDBeShape"
10:00 - 11:00: Practical work. Agnel Praveen Joseph: "Volume matching with PDBeShape"
11:30 – 13:00: Practical work. Students revise model fitting, building and refinement for their structures.

Appendix 5: Cryo-EM Workshop at International Symposium on Grids and Clouds 2017

Course Programme

Monday 6 March 2017

08:30-09:00 Registration

09:00-09:45 General introduction to biological cryoEM / Chi-Yu Fu / Scientific talk

09:45-10:30 Three key breakthroughs that enable cryo-EM to become a routine density map generator / Wei-Hau Chang / Scientific talk

10:30-11:00 Coffee break

11:00-12:30 Relion (part 1) / Wei-Hau Chang / Tutorial

12:30-13:30 Lunch

13:30-15:00 Relion (part 2) / Wei-Hau Chang / Tutorial

15:00-15:30 Coffee break

15:30-17:30 Leggion and Appion / Chi-Yu Fu / Tutorial

Tuesday 7 March 2017

11:00-11:45 Image Processing in cryoEM: Open problems and current perspectives / Jose Maria Carazo / Scientific talk

11:45-12:30 Applications of cryo-electron microscopy to understand complex structures / Sunny Wu / Scientific talk

12:30-13:30 Lunch

13:30-15:00 EMAN 2 (part 1) / Sunny Wu / Tutorial

15:00-15:30 Coffee break

15:30-17:30 EMAN 2 (part 2) / Sunny Wu / Tutorial

Wednesday 8 March 2017

09:45-10:30 High-resolution integrative modelling of biomolecular complexes from fuzzy data / Alexandre Bonvin / Scientific talk

10:30-11:15 POWERFIT and DISVIS (part 1) / Alexandre Bonvin / Tutorial

11:15-11:30 Coffee break

11:30-13:00 POWERFIT and DISVIS (part 2) / Alexandre Bonvin / Tutorial

13:00-13:45 Lunch

13:45-15:15 Scipion (part 1) / Jose Maria Carazo / Tutorial

15:15-15:30 Coffee break

15:30-17:15 Scipion (part 2) / Jose Maria Carazo / Tutorial

17:15-17:30 Closing remark/round table discussion

Appendix 6: Lectures at User Training workshops

- Joana Pereira. "Automatic small molecule identification and ligand building with ARP/wARP", CCP4 Study Weekend, Nottingham, United Kingdom, January 2016

- Alexandre Bonvin. "Integrative modelling of biomolecular complexes using HADDOCK2.2". BioExcel workshop, EBI, Hinxton UK, May 3-4, 2016.

- Alexandre Bonvin. "Integrative modelling of biomolecular complexes: High-resolution modelling with blurry 3D images", Summer Workshop by Biochemistry Division of Korean Chemical Society, KAIST, Daejon, South Korea, June 27, 2016.

- Robbie Joosten. "Validation & optimisation: From a solved structure to the final model", CCP4/APS School in Macromolecular Crystallography, Argonne, IL, USA, June 21-29, 2016.

- Gregor Chojnowski. "Automated Protein Model Building with ARP/wARP 7.4", CCP4/APS School in Macromolecular Crystallography, Argonne, IL, USA, June 21-29, 2016.

- Gregor Chojnowski. "ARP/wARP Automatic Building", CCP4 SPring-8 school: From data processing to structure refinement and beyond, Japan January 23 – 27, 2017.

- Alexandre Bonvin. EMBO practical course on "Integrative modelling of biomolecular complexes", Barcelona, Spain, July 4-9, 2016.

- Alexandre Bonvin. EMBO practical course on "Multidimensional NMR in Structural Biology", Joachimsthal, Germany, July 11-15, 2016.

- Wouter Touw (NKI), "Validation-driven model optimisation". X-Ray Methods in Structural Biology, Cold Spring Harbour, NY, USA, October 10-15, 2016.

- Alexandre Bonvin. "High-resolution, integrative modelling of biomolecular complexes from fuzzy data", Helmholtz Training Course on Integrative Structural Biology, Braunschweig, Germany, November 7-11, 2016

- Robbie Joosten. "Model validation and optimisation", DLS-CCP4 Data Collection and Structure Solution Workshop, Oxford, UK, December 13-20, 2016.

- Alexandre Bonvin. "Information-driven modelling of biomolecular interactions". Ecole thématique – Bioinformatics of protein-protein interactions for wet lab scientists. Pasteur Institute, Paris, April 4-7, 2017

- Alexandre Bonvin. "High-resolution, integrative modelling of biomolecular complexes from fuzzy data". ELIXIR workshop on "Computational Approaches to the Study of Protein Interaction and Drug Design", Padova, Italy, April 10-14, 2017