

Deliverable D7.4



Project Title:	World-wide E-infrastructure for structural biology	
Project Acronym:	West-Life	
Grant agreement no.:	675858	
Deliverable title:	A REFMAC web service for EM and NMR	
WP No.	7	
Lead Beneficiary:	CSIC	
WP Title	Joint Research Activity	
Contractual delivery date:	31 April 2018	
Actual delivery date:	7 May 2018	
WP leader:	Jose Maria Carazo	CSIC
Contributing partners:	EMBL	

Deliverable written by: Victor Lamzin

Deliverable D7.4

Contents

A REFMAC web service for EM and NMR	1
1. Executive Summary.....	3
2. Project Objectives	3
3. Detailed report on the deliverable	4
3.1 Introduction	4
3.2 Details	4
3.3 About ARP/wARP	5

Deliverable D7.4

1. Executive Summary

A web service for cryo-EM models refinement was built based on the next-generation ARP/wARP web infrastructure and using REFMAC ["Refinement of Macromolecular Structures by the Maximum-Likelihood method" G.N. Murshudov, A.A.Vagin and E.J.Dodson, (1997) in Acta Cryst. D53, 240-255.].

The new web service is accessible from the web front end, CCP4i interface and the Balbes web server. The entry has been further expanded to accommodate the submissions from the molecular replacement pipelines Morda and MrBump (Figure 3).

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	x	
2	Provide automated pipelines to handle multi-technique datasets in an integrative manner	x	
3	Provide integrated data management for single and multi-technique projects, based on existing e-infrastructure	x	
4	Foster best practices, collaboration and training of end users	x	

Deliverable D7.4

3. Detailed report on the deliverable

3.1 Introduction

Remote access to the REFMAC engine is provided within the next-generation Web Service of the automated model-building suite ARP/wARP. Here we present basic concepts of the updated service and the description of the file formats used for handling input and output data.

This service has now been extended to support cryo Electron Microscopy (cryo-EM). As compared to structure determination by X-ray diffraction, this has different map features, resolution is a local rather than global property of the data, models can be very large, and maps can be damaged by over-sharpening.

3.2 Details

For over 20 years scientists at EMBL-HH provide to tens of thousands users in the community a web service for the model-building suite ARP/wARP, which uses the REFMAC engine for reciprocal space refinement.

Therefore, to assure a robust and reliable provision the REFMAC web service for cryo-EM models refinement was built based on the next-generation ARP/wARP web infrastructure. The new web service is accessible from the web front end, CCP4i interface and the Balbes web server. The entry has been further expanded to accommodate the submissions from the molecular replacement pipelines Morda and MrBump (Figure 3). Once submitted the jobs are computed at the EMBL-Hamburg Linux cluster containing ~350 CPU cores. The user can monitor the progress of the submitted jobs via the web interface and is also notified by email once a job has finished or terminated prematurely (Figure 1).

The processing of cryo-electron microscopy data starts with parsing and validation of the user-provided density map, primary sequence, and a model and restraints (if any). The web service accepts maps in the MRC and CCP4 formats that are commonly used by electron microscopists. The input model can be provided in the PDB or mmCIF format. The ability to accept the latter by the ARP/wARP web service is essential owing the large sizes of macromolecules studied by cryo-EM. Finally, sequence information that enables model validation and rebuilding (if required) can be accepted in the standard FASTA/PIR formats.

Deliverable D7.4

The workflow diagram presents the main steps of the web service (Figure 2). As a first step, the input data are validated for consistency and any issues are fixed if appropriate and reported to the user. Subsequently, the map and the coordinates are converted to a minimum pseudo-crystallographic asymmetric unit-cell that in some cases may help substantially reduce computation time. The map in the new coordinate frame is then used to calculate structure factors in the MTZ format, which is required by REFMAC. Finally, a REFMAC-based reciprocal-space refinement within a combined iterative ARP/wARP-REFMAC model building is performed depending on the user choice and the characteristics of the data. At the last step, the refined model is placed back to the original reference frame and provided to the user in the PDB file format.

This web service now accepts ARIA identities for login.

3.3 About ARP/wARP

The current version of the ARP/wARP suite and the corresponding web-services are available at

<http://www.embl-hamburg.de/ARP/>

or

<http://www.arp-warp.org/>

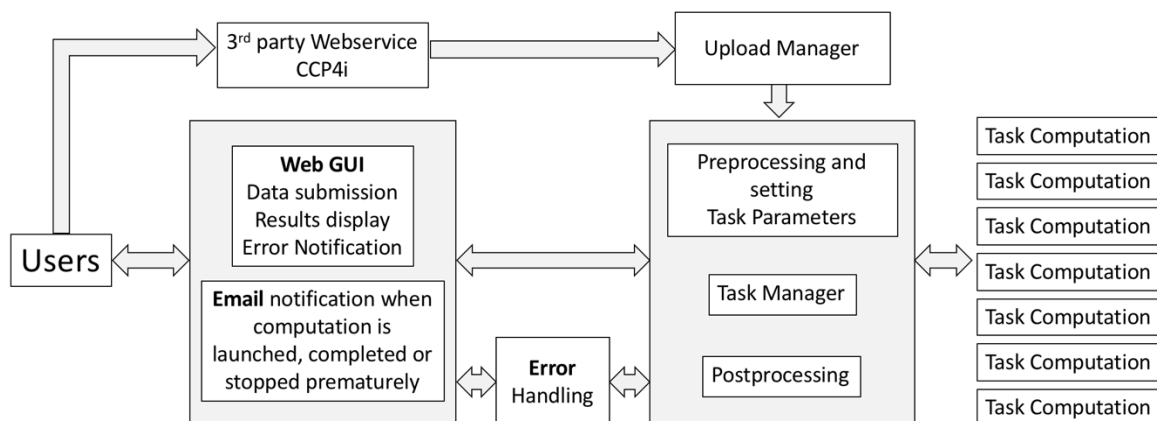


Figure 1. Workflow of the 2nd generation ARP/wARP web service

Deliverable D7.4

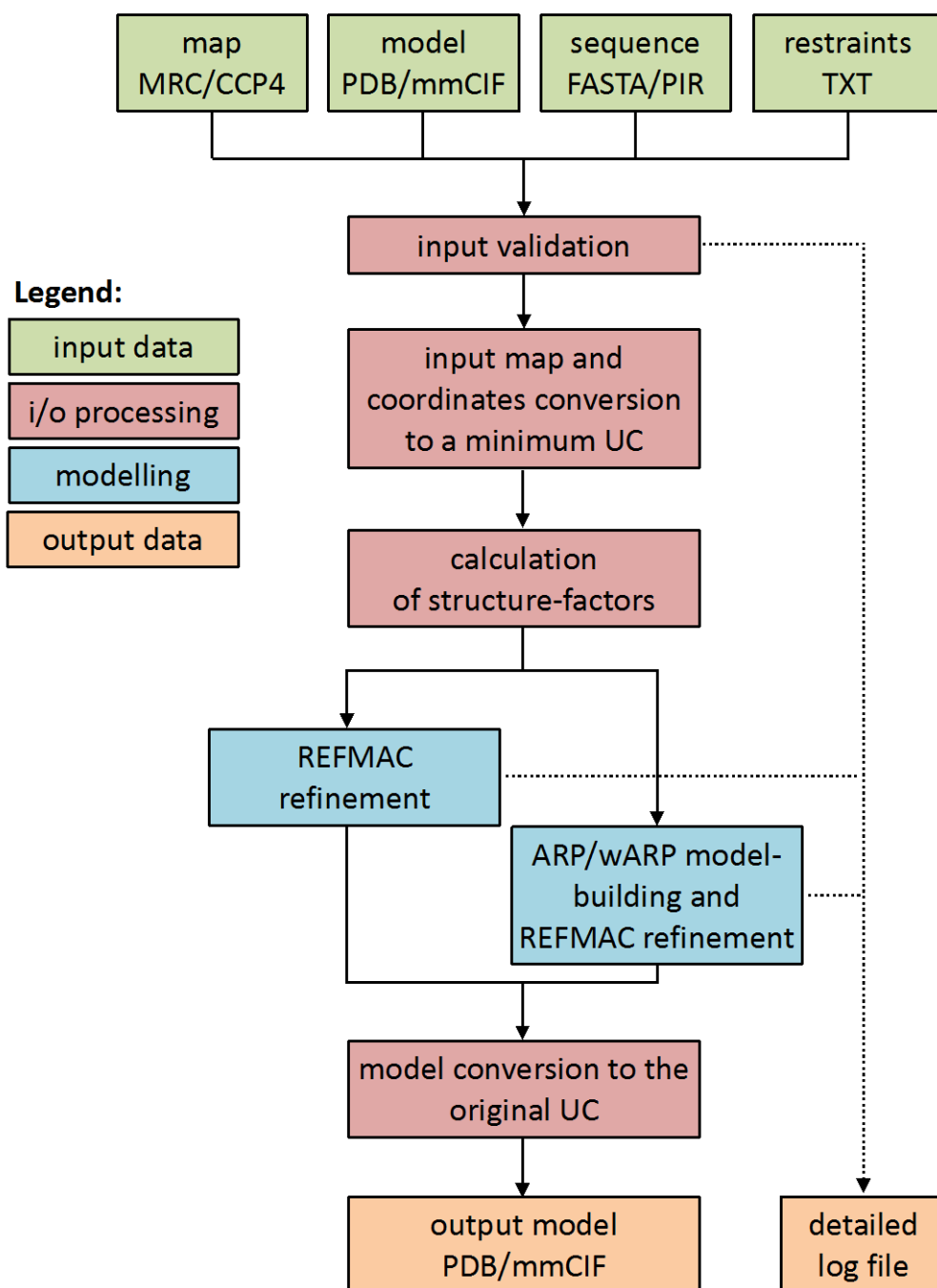


Figure 2. Workflow of the ARP/wARP module for cryo-electron microscopy map interpretation.

Deliverable D7.4

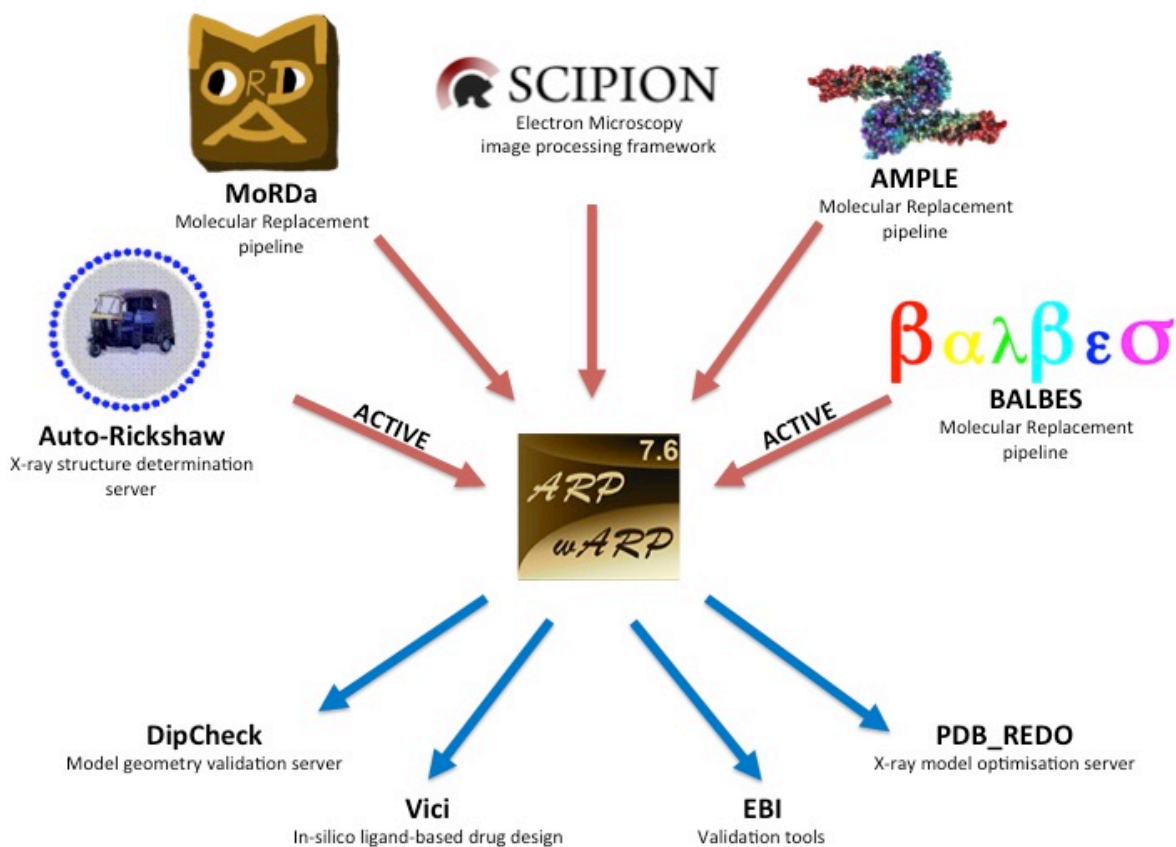


Figure 3. Processing pipelines