ADFExport for OpenLab

ADF Content Description

About This Document

This document provides detailed information on the ADF (Allotrope Data Format) file content created by revision 1.4 of Agilent's ADFExport for OpenLab. It includes a description of the different data sections within an ADF file: ADF data cubes, ADF data description, and the ADF data package.

This document is intended for developers to extract information from ADF files. To read out this information you need to be familiar with the Allotrope Foundation ADF API. For members of the Allotrope Foundation the ADF API SDK is available from the Allotrope Foundation website https://www.allotrope.org.

The ADF files created with the current ADFExport version are based on the Allotrope Framework including the following versions:

- Allotrope Data Framework (ADF) 1.5.1.0
- Allotrope Foundation Ontologies (AFO) REC/2021/03
- Allotrope Data Model (ADM) LC-UV REC/2021/04 (Results-Only data model)
- Allotrope Data Model (ADM) SFC REC/2021/04

1   ADF Data Cubes

This chapter describes the content of the ADF data cubes.

2   ADF Data Description

This chapter describes the content of ADF data description.

3   ADF Data Package

This chapter describes the content of the ADF data package.

4   ADF Audit Trail

This chapter describes the available information in the ADF audit trail
.

5   OpenLab CDS: Injection vs. Sequence based export

This chapter describes the content of an ADF file depending on the OpenLab CDS and OpenLab ChemStation workflow it was created with.

6   Review Content of an ADF file

This chapter describes how to review the content of an ADF file.

7   Appendix

This chapter contains a detailed list of all available parameters in the data description and changes compared to earlier ADFExport revisions.

1 ADF Data Cubes

Data cubes are universal data containers which are used to store analytical raw data. For each chromatogram and UV spectrum in the CDS data, a data cube item with one of the following labels is created (see also "ADF File: Injection vs. Sequence Context" on page 10).

- Chromatogram: adf://dc/<signal name>/<ID>
  where for OpenLab CDS and ChemStation the ID is <signal ID>, (e. g. adf://dc/DAD1A/1600062f-2e19-4f88-8795-766a2b7091b9.

- Spectrum: adf://dc/<signal name>/<ID>
  where for OpenLab CDS and ChemStation the ID is < injection ID>, (e. g. adf://dc/DAD1I/473f2829-ee12-4efe-9f83-046fb7fe60e6.

NOTE Characters allowed for the data cube label are: 0-9, a-z, A-Z, dot, hyphen, and underscore. Empty characters (for example, in the data file name) are replaced by underscores. All other characters are ignored.

Each data cube item is described by information stored in the ADF data description. This description contains axes (absorbance, acquisition time, wavelength) and raw data description/identifier information. For more details, refer to "ADF Data Description" on page 4.

NOTE Instrument signals such as pump pressure or temperature are not exported to ADF data cubes, yet. They are only part of the original files stored in the ADF data package.

If blank subtraction or smoothing is applied during reprocessing the data, the transformed signals are stored instead of the raw data.

2 ADF Data Description

The data description is based on an Allotrope Foundation LC-UV (Results-Only) and SFC semantic model using the W3C standardized RDF format to store an initial set of metadata about the single run or the entire sequence as data graph. The metadata include information about the HPLC/SFC run, such as the system (HPLC/SFC instrument system, including module and column information), the submitter, the sample, the injection, and the sequence. Furthermore, information on chromatograms and UV spectra as well as acquisition and processing method names and peak results are stored.

A schematic representation of the data graph is shown in Figure 1. Depending on the data only parts of the graph can be available. For example, if data are acquired but not further processed only acquisition related parameters are stored without any information on data processing and integration including peak details.
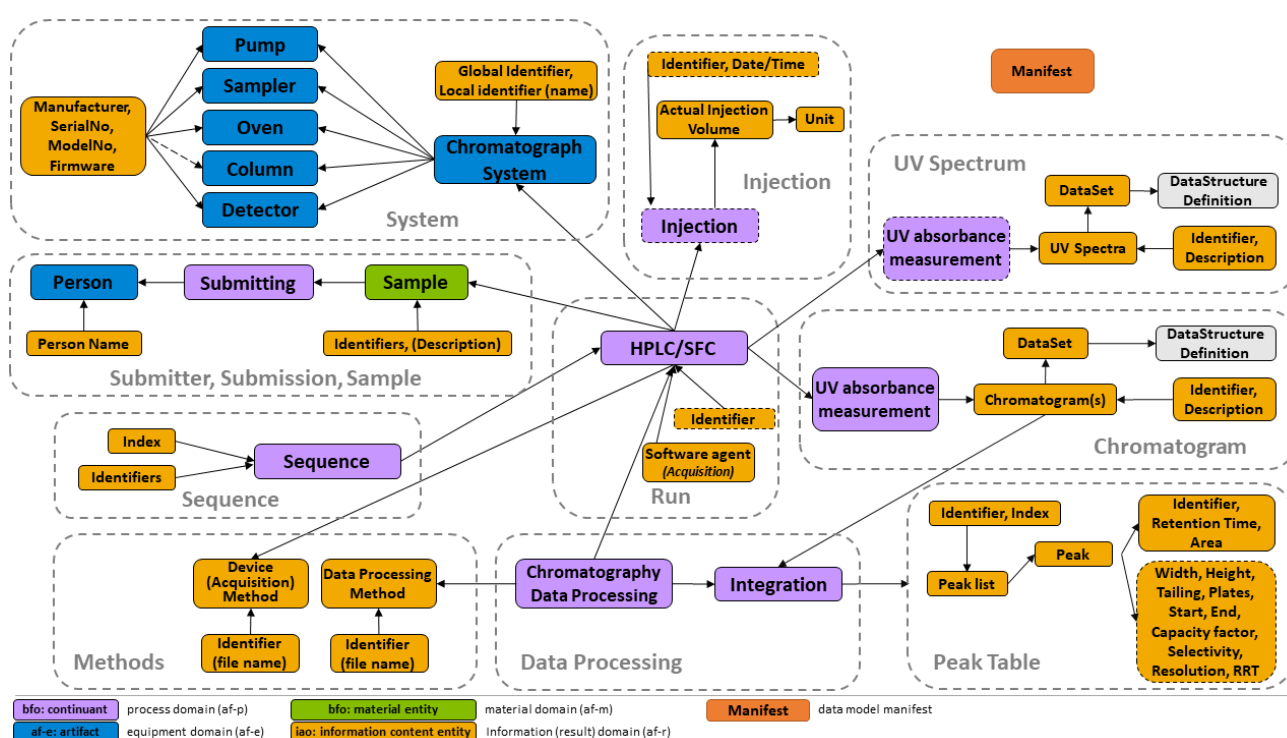

Figure 1   Schematic representation of the metadata which is available in the Data Description of a sequence ADF file.

The data cube structure for chromatograms and spectra is written into the ADF data description and linked to the respective data cube item. For more details, refer to "Data Cube Structure for Chromatograms" and "Data Cube Structure for UV Spectra" on page 5.

A detailed list of all available parameters in the data description is documented in the "Appendix".

NOTE   The data description of an ADF file contains many generated (random) IDs. Due to compression ADF files, originating from the same input data, can have different file sizes. With smaller ADF file sizes this effect becomes more and more negligible.

2.1 Data Cube Structure for Chromatograms

The LC/SFC chromatogram graph structure contains the following values and dimensions:

- Measured values (absorbance in milli absorbance units (mAU))
- Dimension (acquisition time in seconds)

A detailed description of the structure is provided in Table 2.

NOTE The chromatogram data cube structure is written once to the data description of ADF, independent of the number of chromatograms.

OpenLab CDS provides UV chromatogram raw data time values in the time unit milliseconds. ADFExport converts them to seconds, as specified by the Allotrope foundation . OpenLab ChemStation provides these values already in seconds. Therefor no conversion of the ChemStation acquisition time is required.

2.2 Data Cube Structure for UV Spectra

The UV spectra graph structure contains the following values and dimensions:

- Measured values (absorbance in milli absorbance units (mAU))
- Two dimensions (acquisition time in seconds and wavelength in nanometer)

A detailed description of the structure is provided in Figure 3.

NOTE The UV spectra data cube structure is written once to the data description of ADF, independent of the number of spectra.

OpenLab CDS provides UV spectra raw data time values in the time unit milliseconds. ADFExport converts them to seconds, as specified by the Allotrope foundation.

Chromatogram Data Cube Structure Documentation
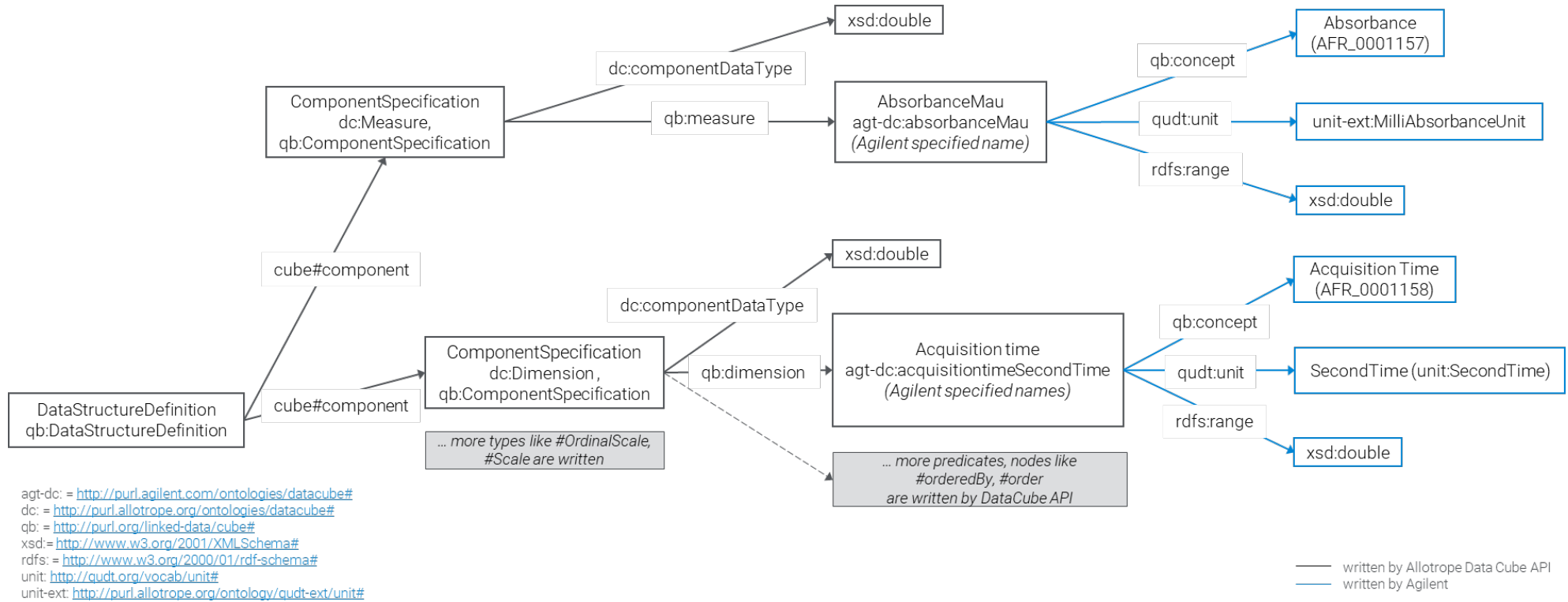
Figure 2 Chromatogram structure (absorbance / acquisition time) written into ADF data description.

UV Spectra Data Cube Structure Documentation

agt-dc: = http://purl.agilent.com/ontologies/datacube#
dc: = http://purl.allotrope.org/ontologies/datacube#
qb: = http://purl.org/linked-data/cube#
xsd:= http://www.w3.org/2001/XMLSchema#
rdfs: = http://www.w3.org/2000/01/rdf-schema#
unit: http://qudt.org/vocab/unit#
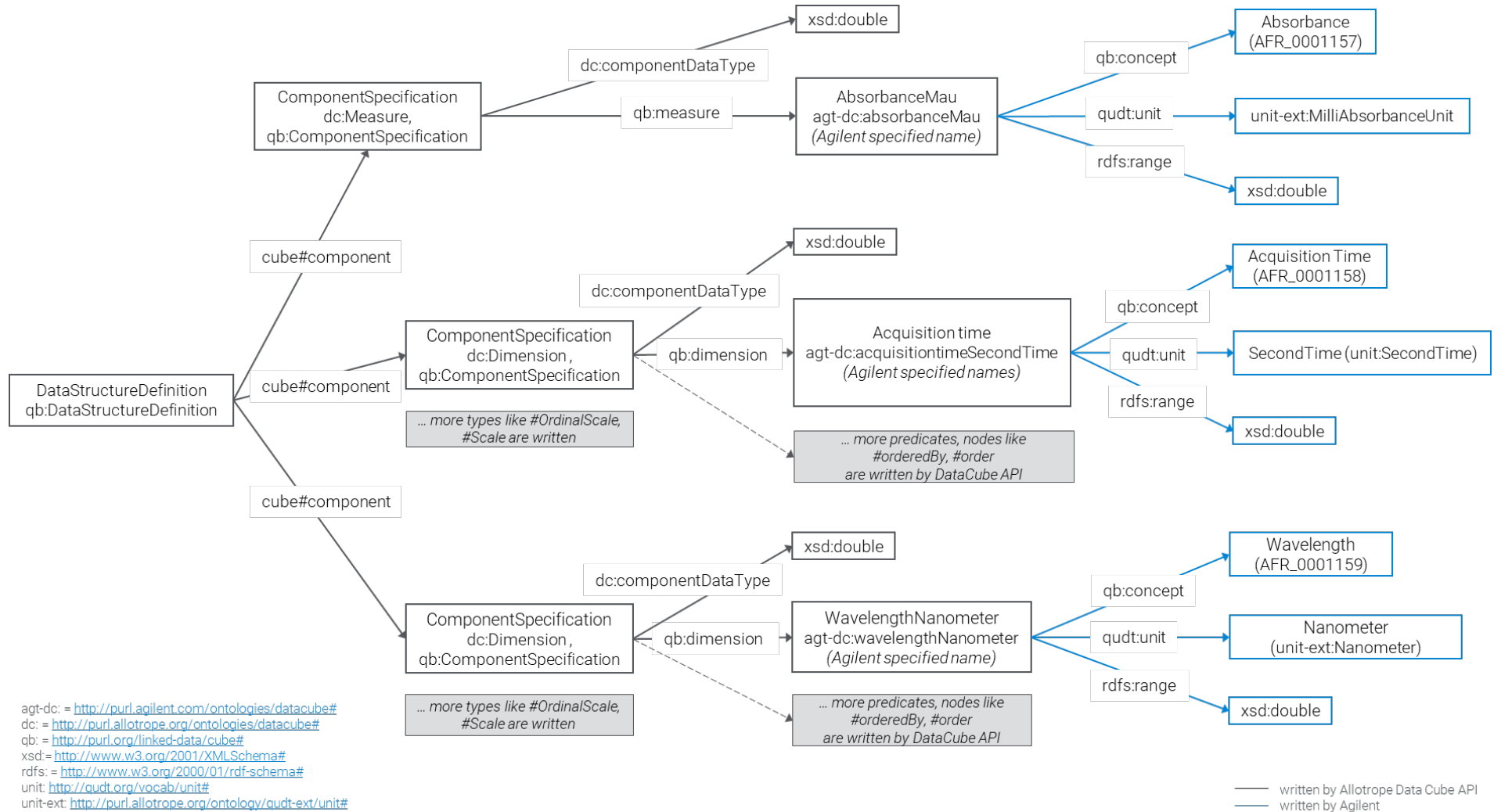unit-ext: http://purl.allotrope.org/ontology/qudt-ext/unit#

Figure 3   UV spectrum structure (absorbance / acquisition time / wavelength) written into ADF data description.

3 ADF Data Package for OpenLab CDS

The data package contains a subfolder **OpenLab CDS** with the original CDS single sample or sequence result folder (*.sirslt or *.rslt). It includes the following files: *.dx (rawdata), *.acaml (metadata), *.amx (acquisition method), *.pmx (data processing method, if available) and *.sqx (sequence). If you export manually in Data Analysis, *.rx files are also exported.

If you export data that has been acquired with ChemStation, the original ChemStation files are exported as well.

NOTE	The *.acaml file is stored in the version of the time of export.

Following characters used for single sample or sequence result folder names will be exchanged by following % encodings:] %5D, [%5B, ^ %5E, ` %60, } %7D, { %7B,~ %7E.

4 ADF Audit Trail

ADF Audit Trail is a functionality provided by the Allotrope Framework and is activated at the end of ADF file generation. After activation all modifications to the ADF file are tracked and grouped into "audit records". The initial audit trail section contains minimal information about the creation of the file, such as date/time and agent (user ID) information.

NOTE To read out the ADF Audit Trail information use an Allotrope Foundation provided tool or create a *.ttl dump file of the ADF file (e.g. to get application name and version information of the tool the ADF file was created with).
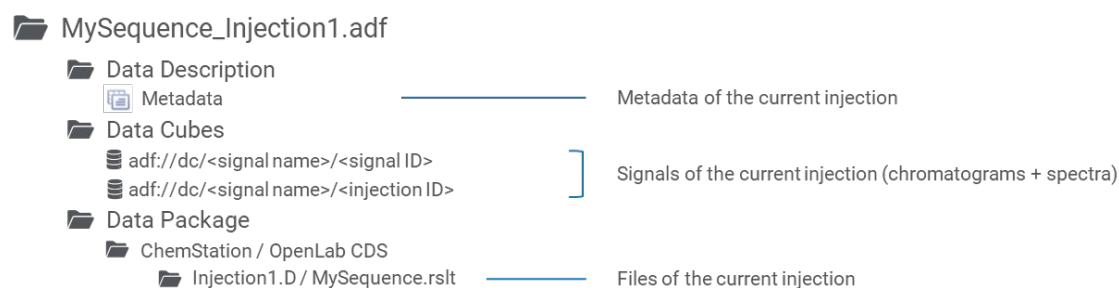
5 Injection vs. Sequence based export

There are different ways to export single sample or sequence data to ADF format. In a sequence injection-by-injection export only the information of a single injection is stored in an ADF file, which results in multiple ADF files for a sequence. Similar for single sample runs (for details, see "Injection based export"). In case of a sequence-based export, the whole sequence information is stored in a single ADF file (for more detail, see "Sequence based export").

5.1 Injection based export

An ADF file is created for the injection of a single run or for each injection of a sequence. A sequence with multiple injections exported therefore results in multiple ADF files, and no sequence context is stored in this case.

An export can be achieved via post-processing plugin either automatically at the end of data acquisition or during reprocessing in OpenLab Data Analysis. For a single sample, one ADF file named SiSa_<result name> is created. In case of a sequence one ADF file is created for each injection of the sequence. The filenames follow the structure: *<sequence name>_<injection file name>.adf*.

Figure 4 Exemplary data containers of an injection ADF file.

5.2 Sequence based export

The entire result set can be exported to a single ADF file. A sequence with multiple injections is exported to one ADF file, and the sequence context is kept.

A single sample or sequence result set (*.sirslt or *.rslt) can be exported manually to ADF in Data Analysis via the "Export ADF" button. The entire result set is stored in one single ADF file.

Review Content of an ADF File

Figure 5   Exemplary data containers of a sequence ADF file.

📂 MySequence.adf
  📂 Data Description
      🖼 Metadata ————————————— Sequence metadata
  📂 Data Cubes
      ⛁ adf://dc/<data file name>/<signal name>/<signal ID> ⎤
      ⛁ adf://dc/< data file name>/<signal name>/<injection ID> ⎦ Signals of injection 1 (chromatograms + spectra)
      ⛁ adf://dc/<data file name>/<signal name>/<signal ID> ⎤
      ⛁ adf://dc/< data file name>/<signal name>/<injection ID> ⎦ Signals of injection 2 (chromatograms + spectra)
          ⋮
  📂 Data Package
      📂 ChemStation / OpenLab CDS
          📂 MySequence / MySequence.rslt ————————— Sequence result folder
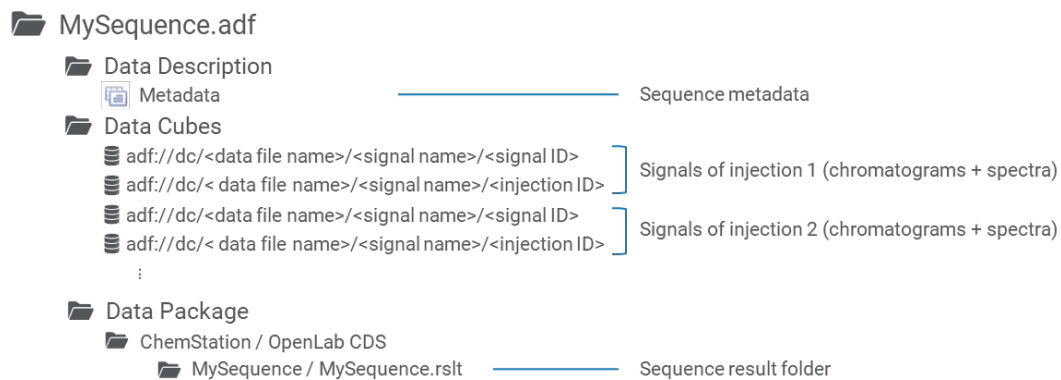
6 Review Content of an ADF File

To browse the content of an ADF file, use the ADF explorer or the ADFSee tool, both available from the Allotrope Foundation (https://www.allotrope.org).
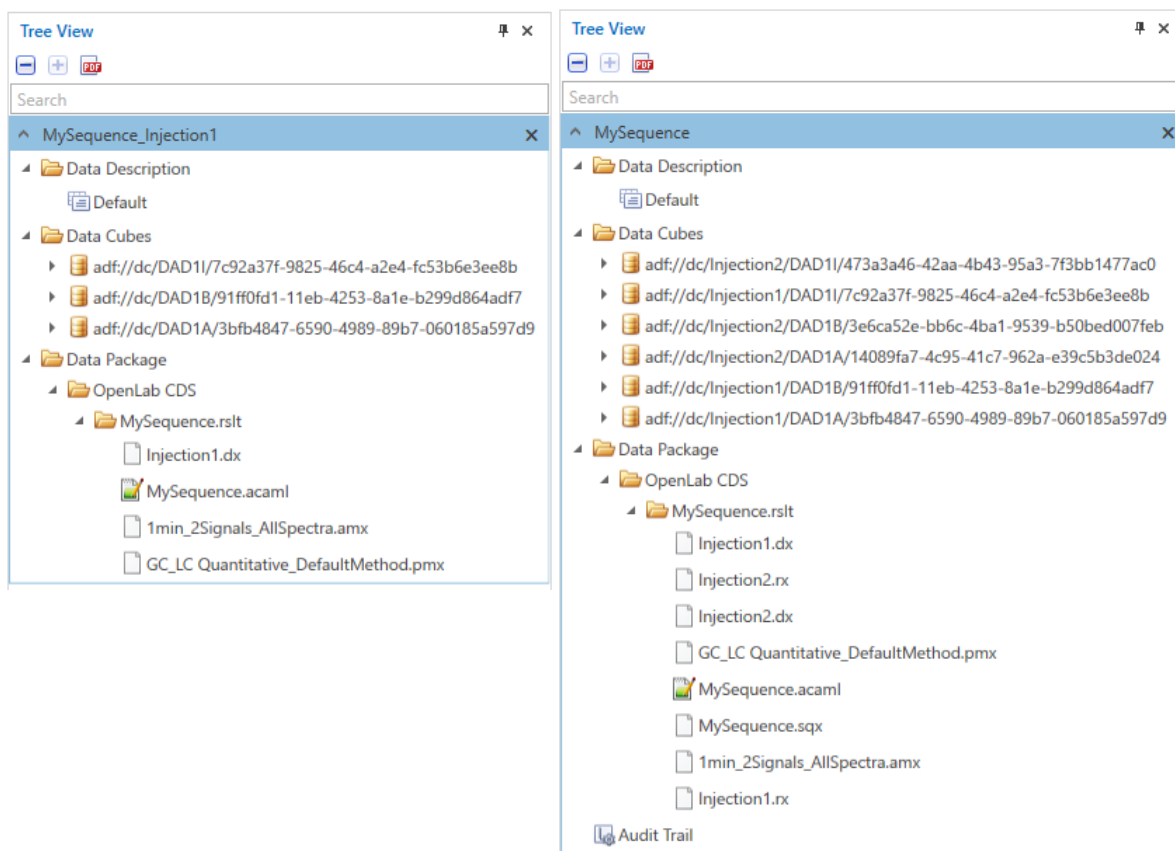
Figure 6   Tree View of ADF explorer, showing a single injection and a sequence ADF file created with OpenLab CDS.

7 Appendix

7.1 Mapping Allotrope Foundation terms to chromatography data system values

Table 1   Mapping of Allotrope to OpenLab CDS parameters.

Allotrope Data Model			OpenLab CDS		
Property name	Identifier	Unit	Name	Unit	Acaml
Run					
Global identifier	AFR_0000918	-	Injection ID	-	<Injections><MeasData><id>
Instrument System					
Written name	IAO_0000590	-	Instrument name	-	<Instrument><Name>
Global identifier	AFR_0000918	-	Instrument ID	-	<Instrument><id>
Instrument System - Module					
Manufacturer	AFR_0001258	-	Vendor	-	<Instrument><Module><Manufacturer>
Written name	IAO_0000590	-	Name	-	<Instrument><Module><Name>
Equipment serial number	AFR_0001119	-	Serial Number	-	<Instrument><Module><SerialNo>
Model number	IAO_0000017	-	Part Number	-	<Instrument><Module><PartNo>
Firmware version	AFR_0001259	-	Firmware Revision	-	<Instrument><Module><FirmwareRevision>
Instrument System - Column					
Equipment serial number	AFR_0001119	-	Serial number	-	<SeparationMedium><Type><Column><SerialNo>
Model number	IAO_0000017	-	Product number	-	<SeparationMedium><Type><Column><ProdNo>
Written name	IAO_0000590	-	Name	-	<SeparationMedium><Name>
Manufacturer	AFR_0001258	-	-	-	-
Sample / Submitter					
Sample - written name	IAO_0000590	-	Sample name	-	<Samples><Setup><IdentParam><Name>
Description	AFR_0000922	-	Sample description	-	<Samples><Setup><IdentParam><Description>
Global identifier	AFR_0000918	-	Sample ID	-	<Samples><Setup><Id>
Submitter – written name	IAO_0000590	-	Acq. Operator	-	<Samples><Setup><Info><CreatedBy><Username>
Injection					
Global identifier	AFR_0000918	-	Injection ID	-	<Injections><MeasData><id>
Scalar timestamp datum	AFR_0001512	-	Injection Time	-	<Injections><MeasData><Info><CreatedDate>

Review Content of an ADF File

Table 1   Mapping of Allotrope to OpenLab CDS parameters.

Allotrope Data Model			OpenLab CDS		
Property name	Identifier	Unit	Name	Unit	Acaml
Autosampler injection volume setting	AFR_0001267	Microliter	Injection Volume	Micro liter	<Injections><MeasData><InjectionVolume>
Sequence					
Global identifier	AFR_0000918	-	Internal parameter	-	<SampleContexts><Setup><id>
Written name	IAO_0000590	-	Sequence name	-	<SampleContexts><Setup><IdentParam><Name>
Index	AFR_0000928	-	Sequence Line	-	*(Index is defined by the acquisition order execution. First ordered by <Injection><Parent sample><OrderNumber>, then by <Injection><Replicate number>)*
Chromatogram					
Local identifier	AFR_0000919	-	Signal name	-	<Signal><Name>
Description	AFR_0000922	-	Signal description	-	<Signal><Description>
UV Spectrum					
Local identifier	AFR_0000919	-	Signal name	-	<Signal><Name>
Description	AFR_0000922	-	Signal description	-	<Signal><Description>
Acquisition Method					
Global identifier	AFR_0000918	-	Method ID	-	<Injections><MeasData><AcqParam><Method_ID>
Local identifier	AFR_0000919	-	Method name	-	<Method><BinaryData><DataItem><Name>
Processing Method					
Global identifier	AFR_0000918	-	Method ID	-	<Injections><Result><DAParam><Method_ID>
Local identifier	AFR_0000919	-	Method name	-	<Method><BinaryData><DataItem><Name>
Integration Method					
Global identifier	AFR_0000918	-	-	-	*(created and assigned by ADFExport)*
Peak Table – Peak List					
Local identifier	AFR_0000919	-	-	-	*(n/a)*
Index	AFR_0000928	-	-	-	*(set by ADFExport based on ordered peak retention time)*
Peak Table – Peak Details					
Local identifier	AFR_0000919	-	Peak ID	-	<SignalResult><Peak><id>
Retention time	AFR_0001089	Second	Retention Time	sec	<SignalResult><Peak><RetentionTime>
Peak area	AFR_0001073	Milli Absorbance Unit * Second	Peak area	mAu* sec	<SignalResult><Peak><Area>
Peak width at 5% of height	AFR_0001265	Second	Width 5%	sec	<SignalResult><Peak><Width_5Perc>
Peak width at half height	AFR_0001266	Second	Width 50%	sec	<SignalResult><Peak><Width_50Perc>
Peak width at baseline	AFR_0001264	Second	Width tangent	sec	<SignalResult><Peak><WidthTangent>
Peak height	AFR_0000948	Milli Absorbance Unit	Height	mAu	<SignalResult><Peak><Height>

Review Content of an ADF File

Table 1   Mapping of Allotrope to OpenLab CDS parameters.

Allotrope Data Model			OpenLab CDS		
Property name	Identifier	Unit	Name	Unit	Acaml
Asymmetry factor measured at 5% height	AFR_0001237	-	Tailing	-	<SignalResult><Peak><TailingFactor>
Number of theoretical plates by peak width at half height	AFR_0001241	-	Plates statistical (EP, JP)	-	<SignalResult><Peak><TheoreticalPlatesEP>
Number of theoretical plates by tangent method	AFR_0001240	-	Plates statistical (USP)	-	<SignalResult><Peak><TheoreticalPlatesUSP>
Peak start time	AFR_0001178	Second	Start time	sec	<SignalResult><Peak><BeginTime>
Peak value at start	AFR_0001179	Milli Absorbance Unit	Baseline Start + Level Start	mAu	<SignalResult><Peak><BaselineStart> + <SignalResult><Peak><LevelStart>
Peak end time	AFR_0001180	Second	End time	sec	<SignalResult><Peak><EndTime>
Peak value at end	AFR_0001181	Milli Absorbance Unit	Baseline End + Level End	mAu	SignalResult><Peak><BaselineEnd> + <SignalResult><Peak><LevelEnd>
Capacity factor	AFR_0001234	-	Capacity factor	-	<SignalResult><Peak><CapacityFactor>
Peak selectivity	AFR_0001235	-	Selectivity	-	<SignalResult><Peak><Selectivity>
Relative peak height	AFR_0000949	Percent	Heigh%	%	<SignalResult><Peak><HeightPercent>
Chromatographic peak resolution using baseline peak widths	AFR_0001231	-	Resolution USP	-	<SignalResult><Peak><Resolution_USP>
Chromatographic peak resolution using peak width at half-height	AFR_0001232	-	Resolution JP, EP	-	<SignalResult><Peak><Resolution_EP>
Unadjusted relative retention time	AFR_0001175	-	RRT	-	<SignalResult><Peak><RelativeRetentionTime>
Adjusted relative retention time	AFR_0001176	-	RRT EP	-	<SignalResult><Peak><RelativeRetentionTime_EP>
Peak area (relative)	AFR_0001165	Percent	Area%	%	<SignalResult><Peak><AreaPercent>
Acquisition Software Agent					
Local identifier	AFR_0000919	-	Acquisition software	-	<Injections> <MeasData> <AcquisitionSoftware>
Manufacturer	AFR_0001258	-	Acquisiton software vendor	-	(part of <Injections> <MeasData> <AcquisitionSoftware>)
Brand name	AFR_0001680	-	Acquisition software brand name	-	<Injections> <MeasData> <AcquisitionApplication> <AgilentApp> <Name>
Software version	AFR_0001700	-	Acquisition software version	-	<Injections> <MeasData> <AcquisitionApplication> <AgilentApp> <Version>

7.2 ADF content history

ADF Export 1.4 vs 1.3

No updates to ADF file content

ADFExport 1.3 vs. ADFExport 1.2

New content:
- Added manifests:
- LC-UV results only: http://purl.allotrope.org/manifests/lc-uv/REC/2021/04/lc-uv-results-only.manifest
- SFC: http://purl.allotrope.org/manifests/sfc/REC/2021/04/sfc.manifest
- Added "acquisition software agent" information linked to "HPLC / SFC Run"
- Added supercritical-fluid chromatography support: "SFC system" (AFE_0002223) and "supercritical-fluid chromatography" (AFP_0000253)

Units:
- For all Data Cube time axis values and peak detail values the time unit is now "seconds" (could be milliseconds, seconds and minute before).
  Reason: Allotrope foundation specified to use in chromatography data models from Dec 2020 onwards, to always use "seconds".
- Injection volume changed Unit from QudtUnit.CubicMillimeter to QudtExt.Microliter (Remark: CubicMillimeter equals Microliter).
  Reason: Allotrope foundation specified to use the new unit term.

Instrument System:
- "LC System" (AFE_0002197) instead of "high-performance liquid chromatograph" (AFE_0000224).
  Reason: Allotrope corrected term class inconsistency. Backward breaking change.

Chromatograms:
- Subclass "electromagnetic radiation chromatogram" (AFR_0000308) instead of parent class: "chromatogram" (AFR_0000236).
  Reason: Use a more precise term of same class family.

Methods:
- Acquisition method. Subclass "device method" (AFR_0001556) instead of parent class: "plan specification" (AFR_0001501). Reason: Use a more precise term is same class family.
- Data processing method. Subclass "data processing method" (AFR_0002175) instead of parent class: "plan specification" (AFR_0001501).
  Reason: Use a more precise term is same class family.