**Use case 1: semantic model representing experimental protocols**

By Olga Giraldo

Experimental protocols are used primarily in life sciences, protocols provide individual sets of instructions that allow scientists to recreate experiments. They provide instructions for the design and implementation of experiments.

It is here described the kind of information that from experimental protocols could be represented in the PlEx ontology.

**p-plan:Plan**

It is a process description (a workflow). i.e., a set of step-by-step instructions where each Instruction describes an action intended by agent(s) to achieve goal(s). An example of a ***plan*** in life sciences is an ***experiment*.**

**p-plan:Step**

It is part of a p-plan:Plan. Each step has input and output Variable(s), and the output of Step SA can be used as the input of Step SB. An example of a ***step*** in life sciences is an ***experimental protocol***

**p-plan:Variable**

It represents a description of the input or output of the planned Activity (p-plan:Step).

Examples of **Input variables** in experimental protocols:

* list of reagents (chemical compounds, mixtures),
* list of lab instruments (equipment, devices, consumables),
* software,
* sample to be tested (whole organism, anatomical part, bodily fluids, etc)
* Datum or dataset

Examples of **Output variables** in experimental protocols:

* processed sample
* Datum or dataset

**bpmn:manualTask**

An action that is done or executed by an agent. An example of manual task in experimental protocols is a ***protocol instruction***.

**Example of Use for the PlEx Ontology**

The classes are presented in ‘***bold and italic’.*** The properties in ‘*italic’*

The experiment (***p-plan:Plan***) titled “Kethoxal-assisted single-stranded DNA sequencing (KAS-seq) for capturing transcription dynamics and enhancer activity” [1], includes 3 experimental protocols (***p-plan:Step***).

1. KAS-seq with mammalian cell cultures,
2. KAS-seq with mouse liver and Low-input KAS-seq,
3. Low-input KAS-seq

In this experiment, the order of execution of the steps or protocols is not relevant, the 3 protocols tested different samples. Therefore, in this example it is not applicable the use of the property *dul:precedes*.

"p-plan:isStepOfPlan": [

{

"@type":"p-plan:Step",

"@id": "http://plex.org/step\_1",

"rdf:value": "KAS-seq with mammalian cell cultures"

},

{

"@type":"p-plan:Step",

"@id": "http://plex.org/step\_2",

"rdf:value": "KAS-seq with mouse liver"

},

{

"@type":"p-plan:Step",

"@id": "http://plex.org/step\_3",

"rdf:value": "Low-input KAS-seq"

}

],

The protocols include input variables like the sample to be tested (***sp:mammalian cell cultures***, ***ncit:mouse liver***), a set of reagents (e.g. ***bao:DMSO***, DBCO-PEG4-biotin, ***snomedct: Boric acid***, and a set of equipment (e.g. ***cseo:Nanodrop***, ***obi:centrifuge***).

"p-plan:hasInputVar": [

{

"@type": "p-plan:Variable",

"@id": "http://purl.org/net/SMARTprotocol#CulturedMammalianCell",

"rdf:value": "mammalian cell cultures"

},

{

"@type": "p-plan:Variable",

"@id": "http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C22515",

"rdf:value": "mouse liver"

},

{

"@type": "p-plan:Variable",

"@id": "http://www.bioassayontology.org/bao#BAO\_0000118",

"rdf:value": "DMSO"

},

{

"@type": "p-plan:Variable",

"@id": "http://plex.org/var\_1",

"rdf:value": "DBCO-PEG4-biotin"

},

{

"@type": "p-plan:Variable",

"@id": "http://scai.fraunhofer.de/CSEO#CSEO\_00001602",

"rdf:value": "Nanodrop"

},

{

"@type": "p-plan:Variable",

"@id": "http://purl.obolibrary.org/obo/OBI\_0400106",

"rdf:value": "Centrifuge"

},

The output variable of the 3 protocols is a ***ncit:single-stranded DNA.***

"p-plan:hasOutputVar": [

{

"@type": "Variable",

"@id": "http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C111311",

"rdf:value": "single-stranded DNA"

}

]

}

The order of execution of each protocol instruction (***bpmn:manualTask*)** can also be represented by using the PlEx ontology. For instance, the experimental protocol titled “KAS-seq with mammalian cell cultures” includes a total of 14 manual tasks. Five of them are represented here.

"pwo:hasFirstStep": [

{

"@type":"bpmn:ManualTask",

"@id":"http://plex.org/manualTask1\_from\_step1",

"description":"Prepare 500 mM N3-kethoxal stock solution using DMSO. Then prepare the labeling medium by diluting the N3-kethoxal solution into pre-warmed (37 ˚C) cell culture medium to a final concentration of 5 mM. It is critical to pre-warm the medium to facilitate N3-kethoxal dissolution.",

"dul:precedes":{

"@type":"bpmn:ManualTask",

"@id":"http://plex.org/manualTask2\_from\_step1",

"description":"Incubate 1-5 million cells in the labeling medium for 10 min at 37 ˚C. For adhesive cells, apply the labeling medium to the cell culture directly in dishes and incubate cells for 10 min at 37 ˚C. For suspension cells, suspend cells in the labeling medium and incubate them for 10 min at 37 ˚C.",

"dul:precedes":{

"@type":"bpmn:ManualTask",

"@id":"http://plex.org/manualTask3\_from\_step1",

"description":"Harvest cells after the incubation. Isolate total DNA from cells by using PureLink genomic DNA mini kit (Thermo K182001). Elute DNA by using 50 µL 25 mM K3BO3 (pH 7.0).",

"dul:precedes":{

"@type":"bpmn:ManualTask",

"@id":"http://plex.org/manualTask4\_from\_step1",

"description":"Prepare the click reaction mixture as the follows: 2 µg total DNA diluted in 85µL of 25 mM K3BO3 (pH 7.0), 10 µL 10× PBS, 5 µL 20 mM DBCO-PEG4-biotin (DMSO solution, Sigma 760749). Incubate the mixture at 37 ˚C for 1.5 h with shaking at 500 rpm to facilitate the “click” reaction.",

"dul:precedes":{

"@type":"bpmn:ManualTask",

"@id":"http://plex.org/manualTask5\_from\_step1",

"description":"Add 5 µL RNase A (Thermo 12091039) to the reaction mixture. Incubate the mixture at 37 ˚C for 15 min with shaking at 500 rpm.",

**Missing information**

The PlEx ontology cannot represent the following information about experimental protocols:

1. The overall objective of the protocol

The description of the objective should make it possible for a user to decide on the suitability of the protocol for their experimental problem.

In the current example the objective of the protocol is presented in the abstract and introduction section. It is not clearly stated as a single paragraph.

1. Alert messages

Such as critical steps, pause points and execution time complement the description of steps and facilitate the correct execution of a protocol. The goal is to remind or alert the user of a protocol with respect to issues that may arise when executing a step. These messages may cover special tips or hints for performing a step successfully, alternative ways to perform the step, warnings regarding hazardous materials or other safety conditions, time considerations.

In the current example the protocols include pause points (or stop points). For example, see the steps 3, 6 and 12 from the protocol titled “KAS-seq with mammalian cell cultures”.

In addition, the same protocol, in step 14 includes the following tips: “**DNAs can be stored at -20 ˚C at the noted stop points if needed.”**

Finally, estimated time of performance of the procedure is presented in the section time Taken.

**References**

1. Tong Wu, Ruitu Lyu, Chuan He et al. Kethoxal-assisted single-stranded DNA sequencing (KAS-seq) for capturing transcription dynamics and enhancer activity, 01 May 2020, PROTOCOL (Version 2) available at Protocol Exchange [https://doi.org/10.21203/rs.3.pex-835/v2]