
PSM CBOR format documentation

version: 0.1

Olivier Langella

Contents

Introduction	3
psm CBOR format	3

Introduction

PSM CBOR file is designed to record any PSM (peptide spectrum match) in a compact binary file, easy to manipulate, versatile, extendable. This file is used as a stream in any condition, allowing the users to use unix pipes, compression algorithms, network transparency.

This way, from a DDA identification engine search result converted in PSM CBOR, any process can be added :

- Feature computations
- Prediction process (retention times, ion mobility, MS2 prediction...)
- Rescoring
- Filtering

psm CBOR format

It could be something like that :

PSM cbor file format

```
{
  "informations": {
    "executable": "mcql",
    "type": "psm",
    "pappsomspp_version": "0.9.46",
    "sysinfo_machine_hostname": "proteus1",
    "sysinfo_product_name": "Debian GNU/Linux 12 (bookworm)",
    "timestamp": "2025-06-25T10:33:51"
  },
  "parameter_map": {
    "sage": {},
    "xtandem": {},
    "specoms": {}
  },
  "target_fasta_files": ["zea_mays.fasta", "contaminant.fasta"],
  "decoy_fasta_files": ["rev_zea_mays.fasta", "rev_contaminant.fasta"],
  "protein_map": {
    "GRMZM2G083841_P01":
      {
        "description": "",
        "sequence": "",
        "target": true,
        "contaminant": false,
        "eval": {}
      },
    "GRMZM2G083841_P02":
      {
      }
  },
  "sample_list": [
    {
      "name": "tandem2017_nopatch_20120906_balliau_extract_1_A01_urnb-1",
      "identification_file_list": [{
        "name": "/home/langella/data1/tandem/
```

```
tandem2017_nopatch_20120906_balliau_extract_1_A01_urnb-1.xml",
  }],
  "peaklist_file": {
    "name": "tandem2017_nopatch_20120906_balliau_extract_1_A01_urnb-1.mzml"
  },
  "scan_list": [
    {
      "id": {
        "index": 1976
      },
      "precursor": {
        "z": 2,
        "mz": 1120.529471
      },
      "ms2": {
        "rt": 12648.87,
        "mz" : [1,2,3,4],
        "intensity" : [1,2,3,4]
      },
      "psm_list": [
        {
          "proforma": "AQEEM[+15.99491]AQVAK",
          "protein_list" : [
            {
              "accession": "GRMZM2G083841_P01",
              "position": [15,236]
            }
          ],
          "eval": {
            "xtandem" : {
              "evaluate": 0.0011
            },
            "specoms" : {
              "evaluate": 0.0011
            }
          }
        }
      ],
    }
  ],
},
]
```