

Spoms documentation

PeptidOMS-ANR-24-CE45-3296

11/07/2025 - 10/04/2026

Olivier Langella
PAPPSO
2025-2026
technical documentation

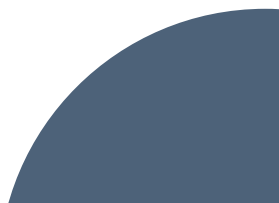


Table of content

1	Introduction	3
2	Source code	3
3	PSM CBOR file workflow	3
4	Command line interface	3
5	JSON parameter file	3
5.1	spectrum section	4
6	CBOR peptidoms eval PSM structure	4
7	Tests for Spoms	5
7.1	Smaller dataset	5
	Bibliography	5

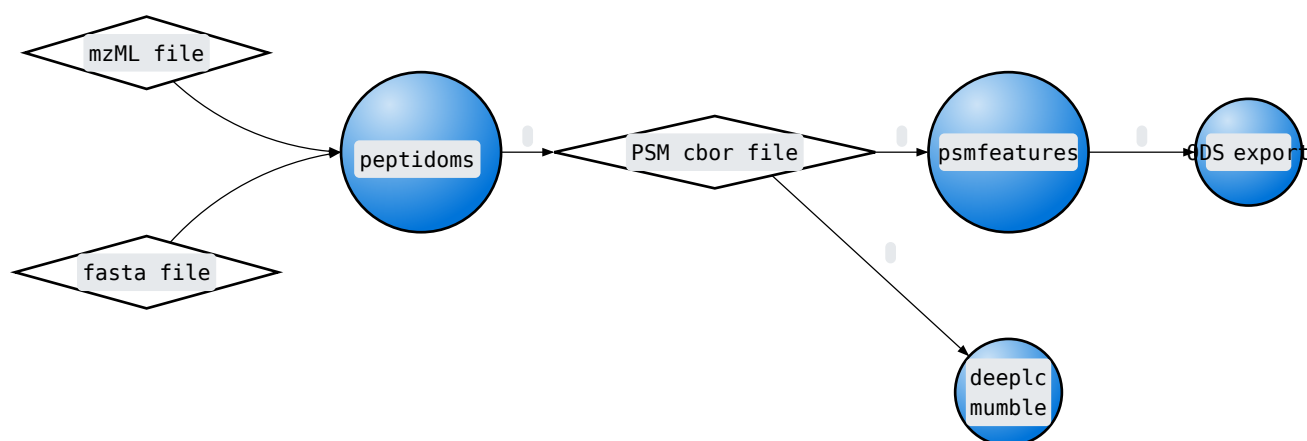
1 Introduction

Spoms aligns MS2 fragmentation spectra directly to protein sequences. Algorithm is described in (Benoist et al., 2025).

2 Source code

Current source code is available at <https://forge.inrae.fr/peptidoms/specpeptidoms>.

3 PSM CBOR file workflow



4 Command line interface

Spoms works on a collection of MS/MS spectra (mzML file typically) and a protein sequence database (FASTA file) and produces a PSM CBOR file as output.

```
spoms -p params.json -m mzdata.mzml -f protein_sequences.fasta -o results.cbor
```

To get a readable file results, use “spoms-export” to produce JSON or ODS files:

```
spoms-export -i results.cbor --ods results.ods
```

```
spoms-export -i results.cbor --json results.json
```

5 JSON parameter file

Spoms parameters are specified in a JSON file

JSON parameter file

```
{
  "spectrum": {
    "fragment_tolerance_unit": "dalton",
    "fragment_tolerance": 0.02,
    "deisotope": true,
    "minimum_mz": 150,
    "n_most_intense": 60
  },
  "alignment": {
    "nb_peaks_min": 30,
    "nb_min_aa_found": 20,
    "min_scenario_score": 10,
    "locations_score_threshold": 0.9,
    "nb_locations_saved": 1,
    "nb_returned_alignments": 2
  },
  "post_treatment": {
    "shut_down_peaks_cleaning": false,
    "shut_down_non_aligned_mass": false,
    "nb_interpretations_saved": 1
  },
  "output": {
    "max_interpretations_per_spectrum": 10
  }
}
```

5.1 spectrum section

fragment_tolerance_unit string “dalton” by default

fragment_tolerance float mass delta precision (0.02 dalton by default)

deisotope boolean if true, process spectrum to merge isotopes into single charge peaks

minimum_mz float minimum m/z ratio to consider peaks in the spectrum

n_most_intense int consider only the N most intense peaks in the spectrum

6 CBOR peptidoms eval PSM structure

Spoms writes results in a PSM CBOR file. For each PSM, Spoms uses the “eval” section to store its scores :

spoms PSM eval structure

```
{
  "proforma": "AIADGSLLDLLR",
  "protein_list": [
    {
      "accession": "GRMZM2G083841_P01",
      "position": [429]
    }
  ],
  "eval": {
    "spoms": {
      "first": {
        "bracket": "[A][I]ADGSLLDLLR",
        "spc": 12,
        "score": 76,
        "nam": 0
      },
      "post": {}
    }
  }
}
```

7 Tests for Spoms

```
./src/spoms -p ../doc/typst/json/spoms_parameters.json -f /gorgone/pappso/moulon/database/Genome_Z_mays_5a.fasta -m /gorgone/pappso/data_extraction_pappso/mzML/20120906_balliau_extract_1_A01_urnb-1.mzML -o /tmp/test.cbor -c 10
```

```
./src/spoms -p ../doc/typst/json/spoms_parameters.json -f /gorgone/pappso/moulon/database/Genome_Z_mays_5a.fasta -m /gorgone/pappso/data_extraction_pappso/mzcbor/20120906_balliau_extract_1_A01_urnb-1.mzcbor -o /tmp/test.cbor -c 10
```

Process for spectrum index 2406 finished. Slowest one is 2287 2290
total 236

7.1 Smaller dataset

```
./src/spoms -p ../doc/typst/json/spoms_parameters_test.json -f ../tests/data/fasta/small_zea_mays.fasta -m ../tests/data/peaklists/peaklist_15046.mgf -o /tmp/test.cbor
```

```
./src/spoms -p ../tests/data/parameters/DATNVGDEGGFAPNIIENK_parameters.json -f ../tests/data/fasta/databankThreeProteinsHuman.fasta -m ../tests/data/peaklists/DATNVGDEGGFAPNIIENK.mgf -o DATNVGDEGGFAPNIIENK.cbor -c 10
```

Bibliography

Benoist, É., Jean, G., Rogniaux, H., Fertin, G., and Tessier, D. (2025). SpecPeptidOMS Directly and Rapidly Aligns Mass Spectra on Whole Proteomes and Identifies Peptides That Are Not Necessarily Tryptic: Implications for Peptidomics. *Journal of Proteome Research* 24, 2159–2172. doi: [10.1021/acs.jproteome.4c00870](https://doi.org/10.1021/acs.jproteome.4c00870)