PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

PSDI Cross Data Search

Open Databases Integration for Materials Design
September 2025

Tom Underwood[1], Mehmet Giritli[1], Alexander Belozerov[1], Aileen Day[2]

[1] *Science Technology Facilities Research Council, UK*
[2] *University of Southampton, UK*

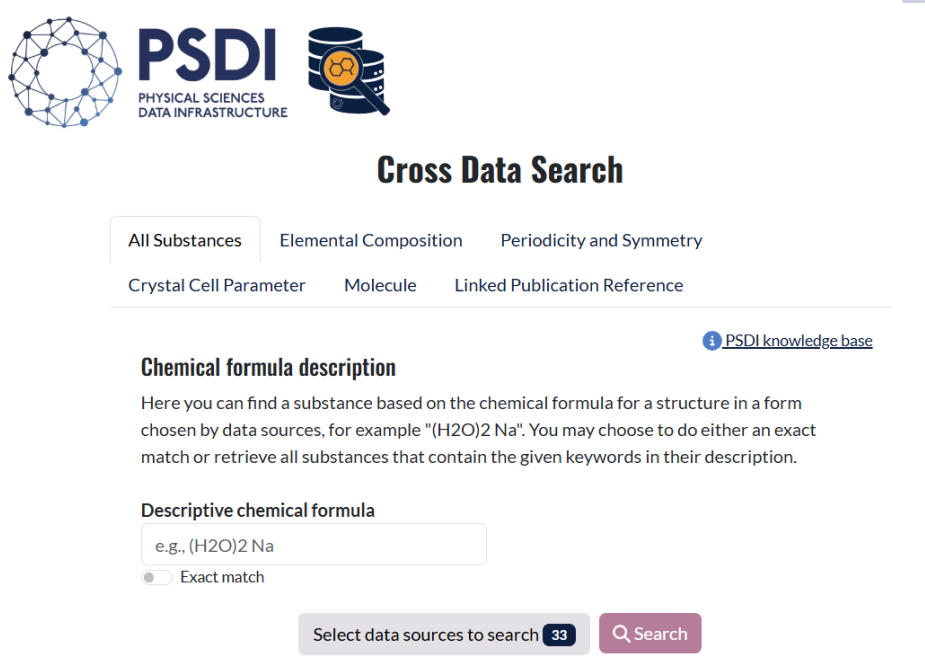
University of Southampton

UKRI Science and Technology Facilities Council
Scientific Computing

UKRI Engineering and Physical Sciences Research Council

PSDI Cross Data Search

- Powered by OPTIMADE (Python tools)!
- Covers many data sources
  - ‘External’ OPTIMADE providers
  - ‘Internal’ PSDI OPTIMADE providers
- Users can search using various approaches:
  - Elemental composition
  - Crystal cell parameters
  - Chemical identifier (SMILES or InChI)
  - Publication reference
  - Advanced search = full OPTIMADE query
- Users can select data sources to query
- Explore custom fields for data sources



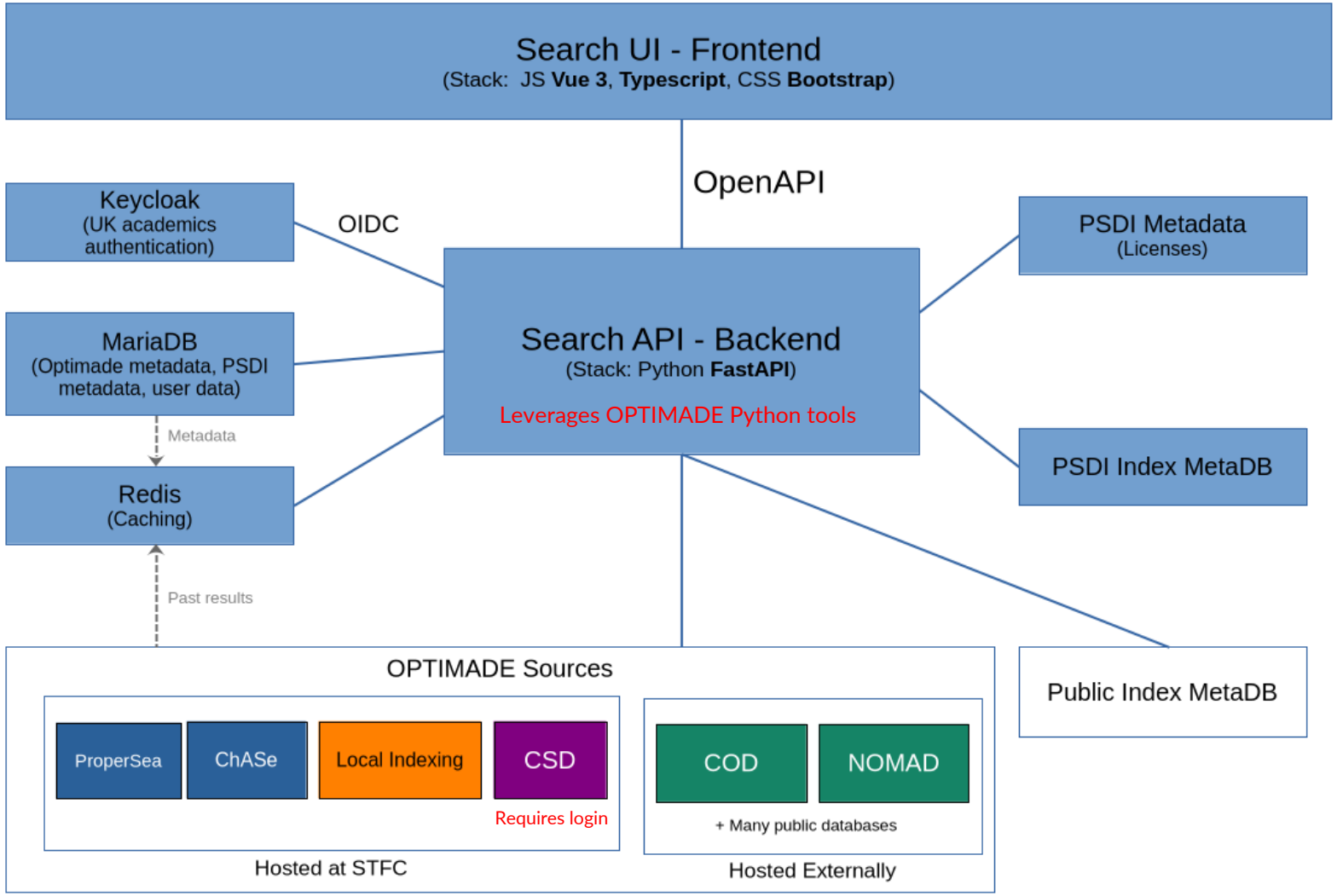
‘Internal’ PSDI OPTIMADE sources

- Propersea
- ChASe
- CCDC
- UK Catalysis Hub Data
- UK Catalysis Hub Publications
- STFC Open Data
- Chemotion

Try it yourself: https://data-search.psdi.ac.uk

PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

Cross Data Search architecture



Cross Data Search Demo

https://data-search.psdi.ac.uk

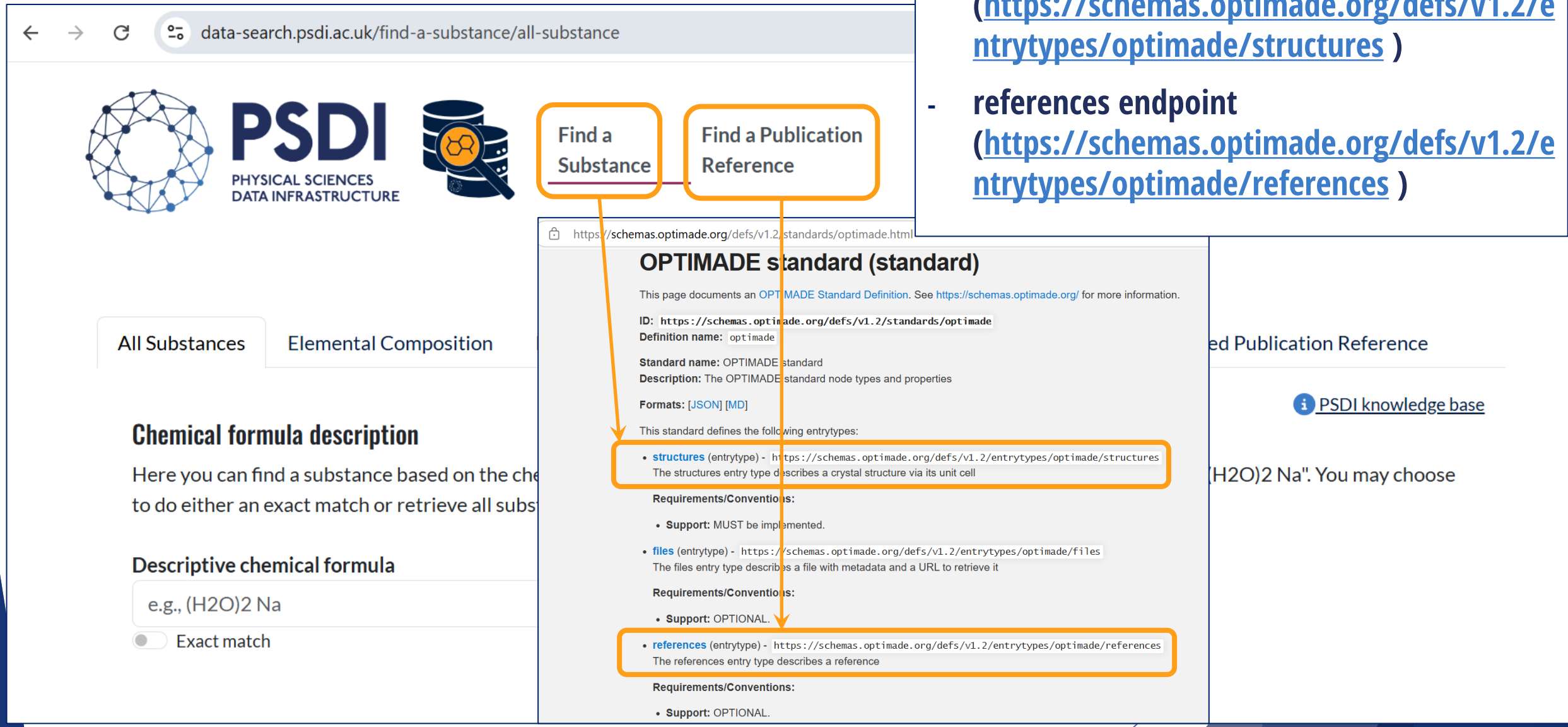
Appendix

Annotated screenshots of Cross Data Search

PSDI Cross Data Se

Align with OPTIMADE properties in:

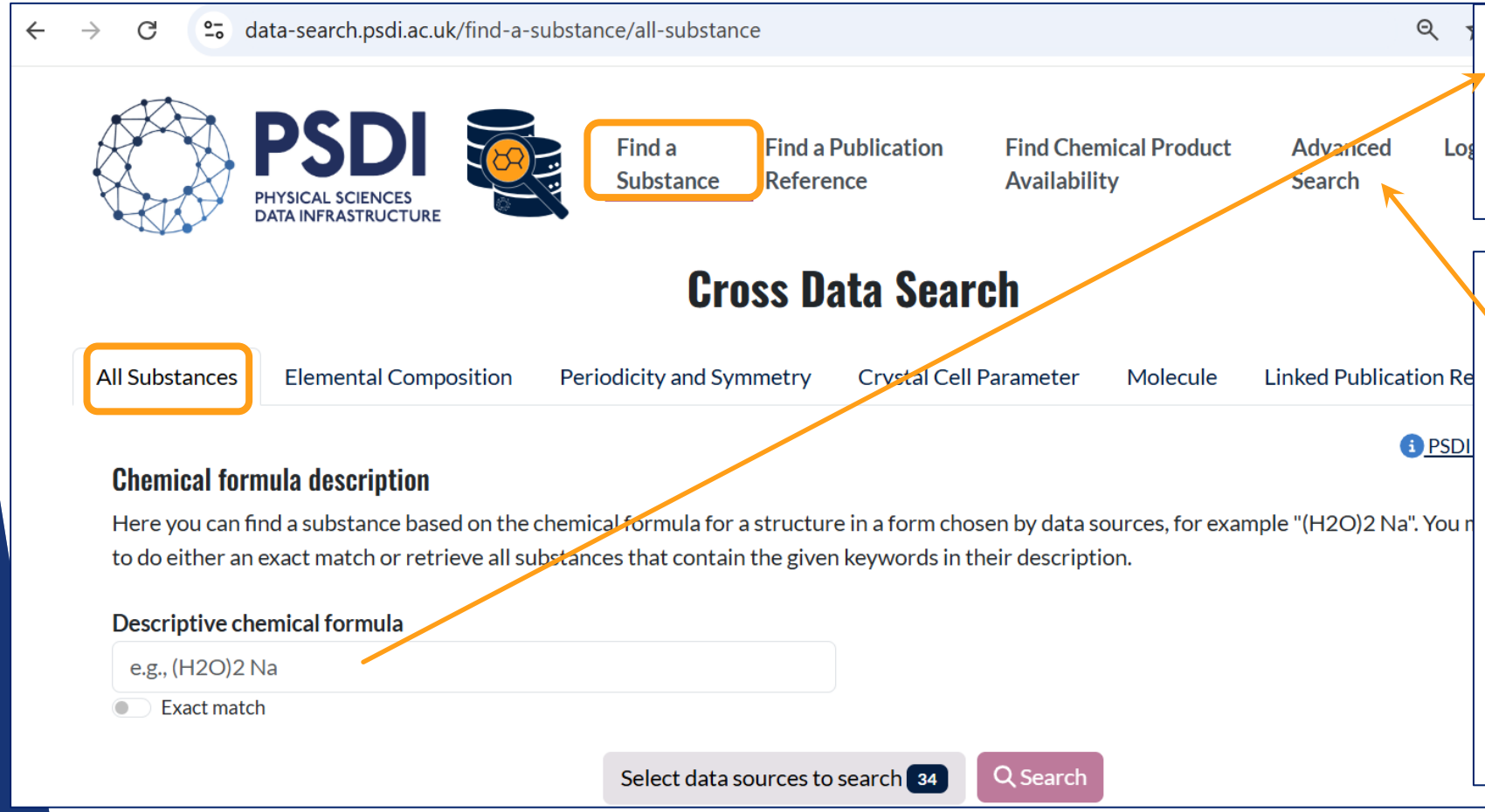
- structures endpoint (https://schemas.optimade.org/defs/v1.2/entrytypes/optimade/structures)
- references endpoint (https://schemas.optimade.org/defs/v1.2/entrytypes/optimade/references)



PSDI – find a substance

What is a substance?

- everything made up of elements



searching on OPTIMADE structure endpoint property:

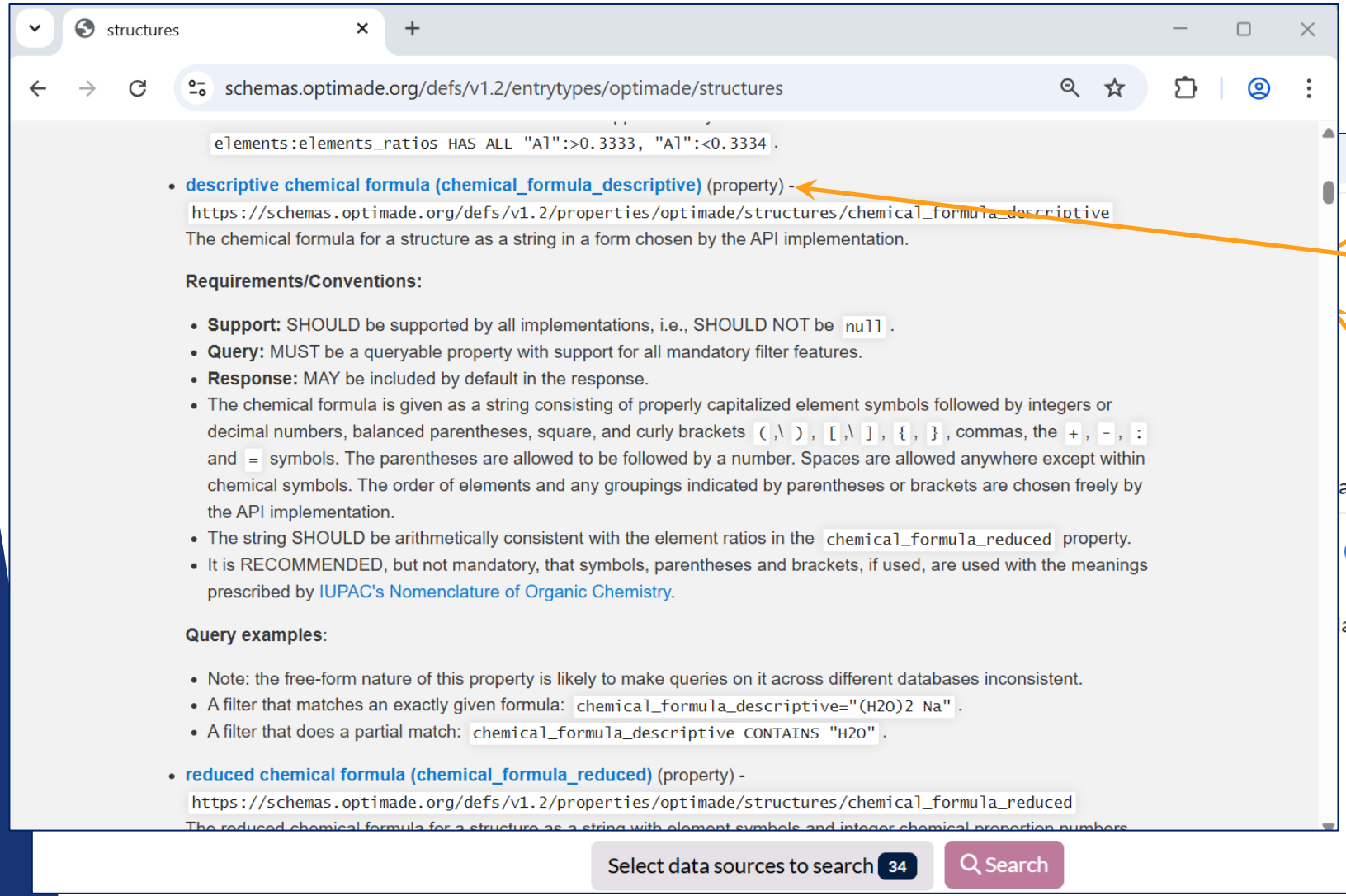
- descriptive chemical formula (chemical_formula_descriptive)

This is the most widely-populated chemical formula property but others are available for searching in the "Advanced Search":

- reduced chemical formula (chemical_formula_reduced)
- Hill chemical formula (chemical_formula_hill)
- anonymous chemical formula (chemical_formula_anonymous)

PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

PSDI – find a substance



OPTIMADE provides comprehensive definition of core properties e.g.

searching on OPTIMADE structure endpoint property:

- descriptive chemical formula (chemical_formula_descriptive)

This is the most widely-populated chemical formula property but others are available for searching in the “Advanced Search”:

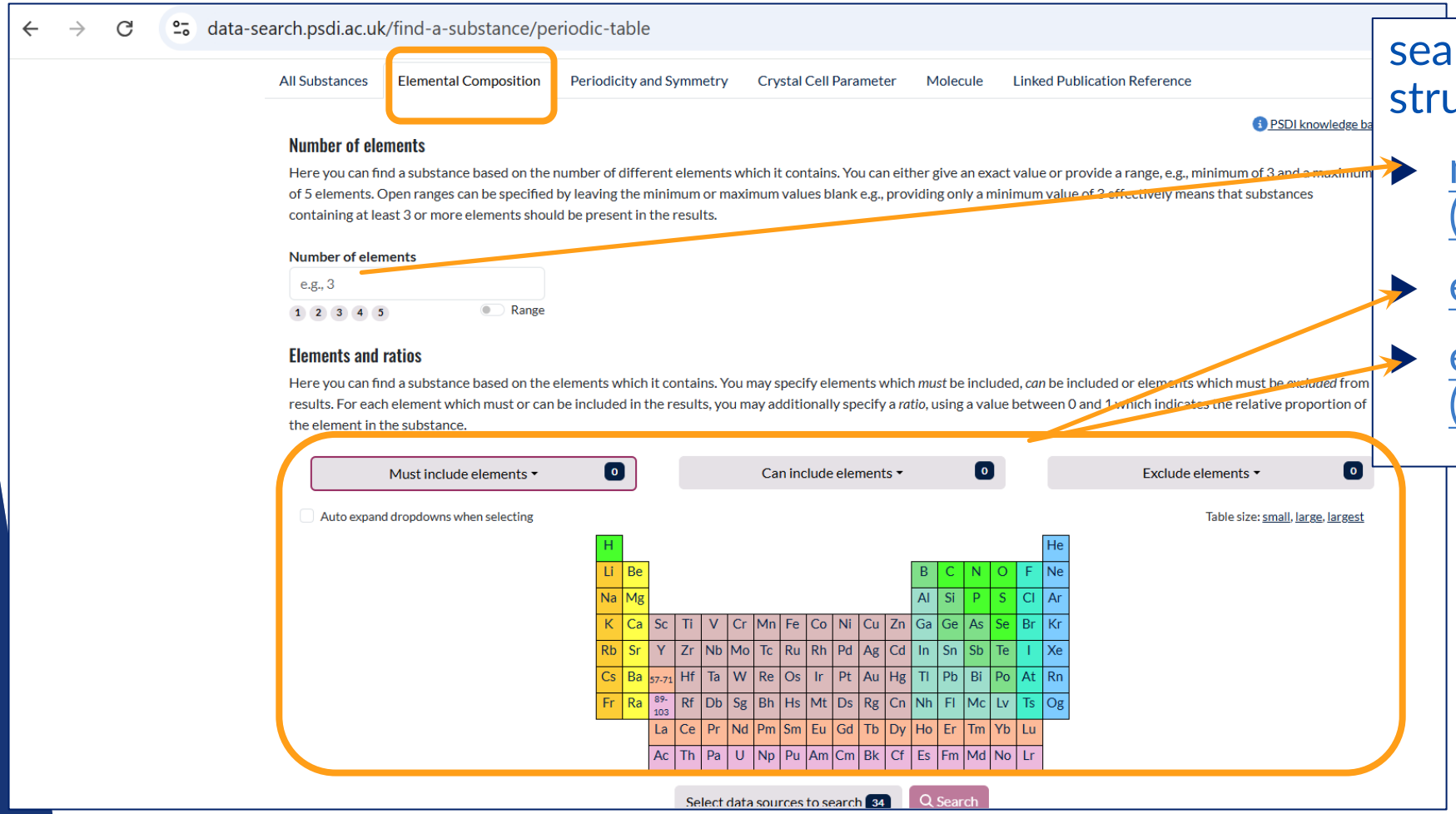
- reduced chemical formula (chemical_formula_reduced)
- Hill chemical formula (chemical_formula_hill)
- anonymous chemical formula (chemical_formula_anonymous)

PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE

PSDI – find a substance

What is a substance?

- everything made up of elements



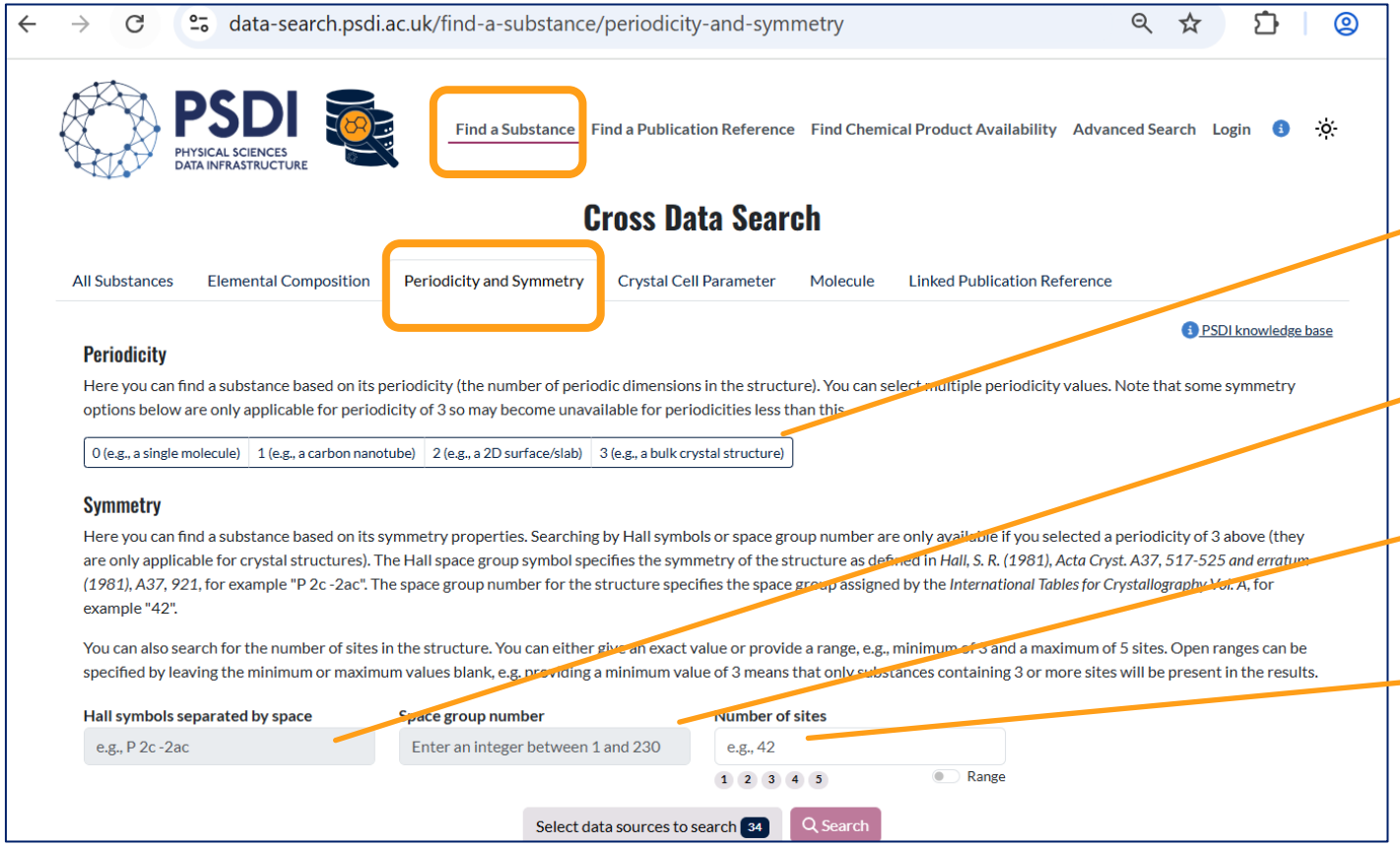
searching on OPTIMADE structure endpoint property:

- number of elements (nelements)
- elements
- elements ratios (elements_ratios)

PSDI – find a substance

If it has perioidicity and symmetry:

- 3-Dimensional (crystal), 2-Dimensional (surface), 1-Dimensional (wire)…
- molecules are zero-dimensional crystal structure in OPTIMADE

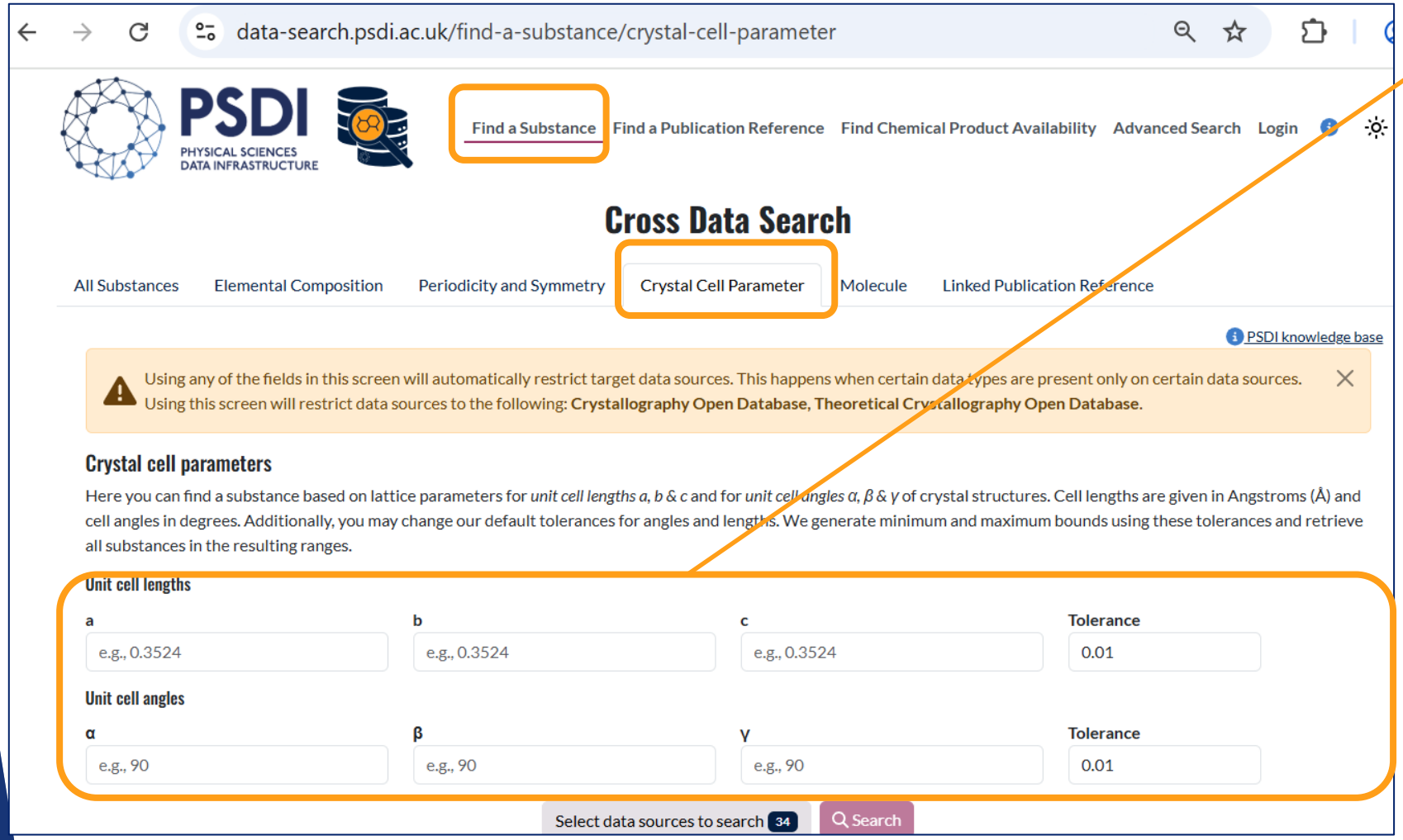


searching on OPTIMADE structure endpoint properties:

- dimension types (dimension_types)
- Hall space group symbol (space_group_symbol_hall)
- space group IT number (space_group_it_number)
- number of sites (property)

PSDI – find a substance

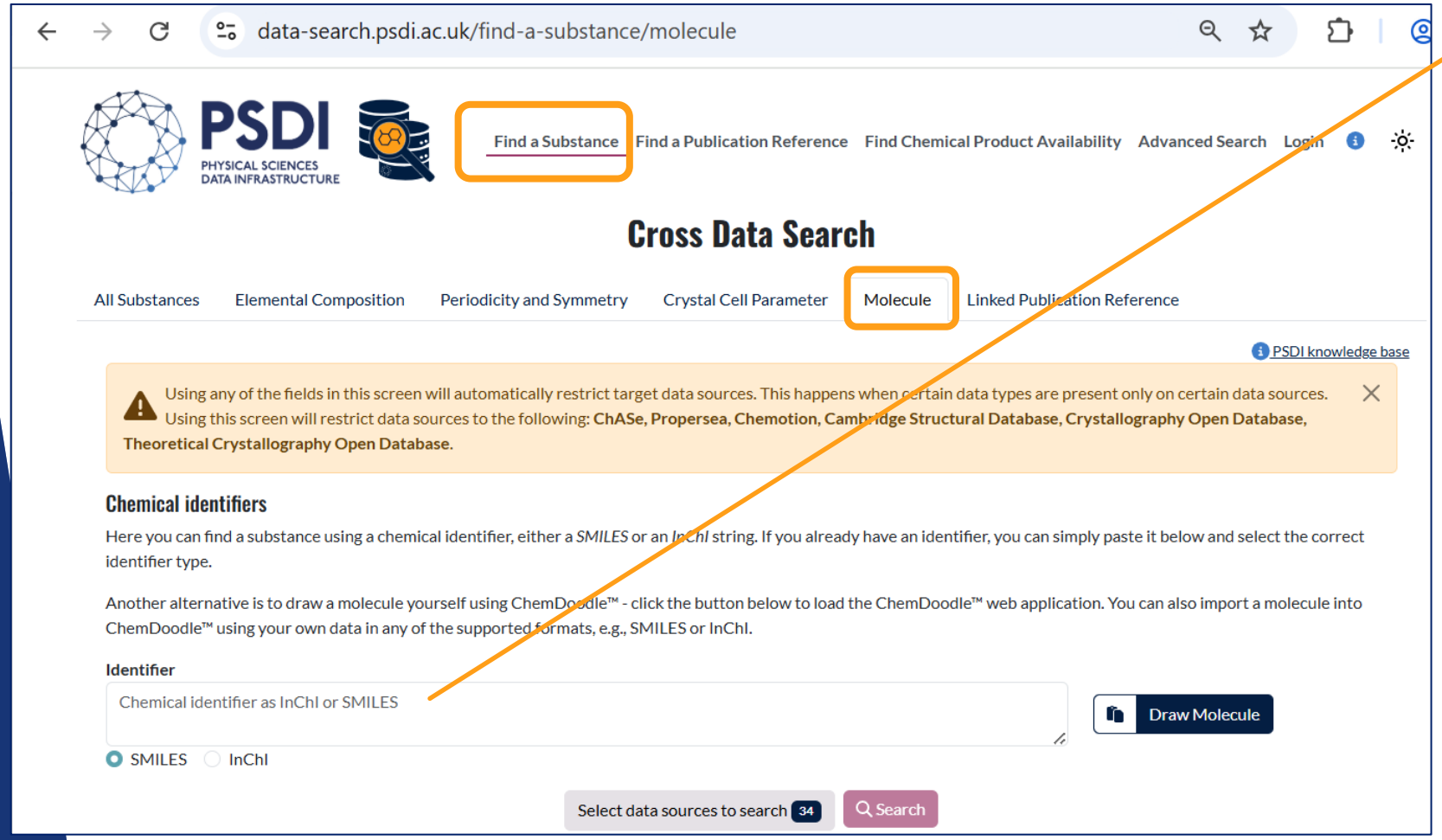
- If it is a crystal structure



- searching on custom properties for cell parameters: a, b, c, alpha, beta, gamma
- soon to be defined in PSDI namespace
- (note that OPTIMADE captures the alternative lattice vectors (lattice_vectors) which we do not use but can be searched on in Advanced Search)

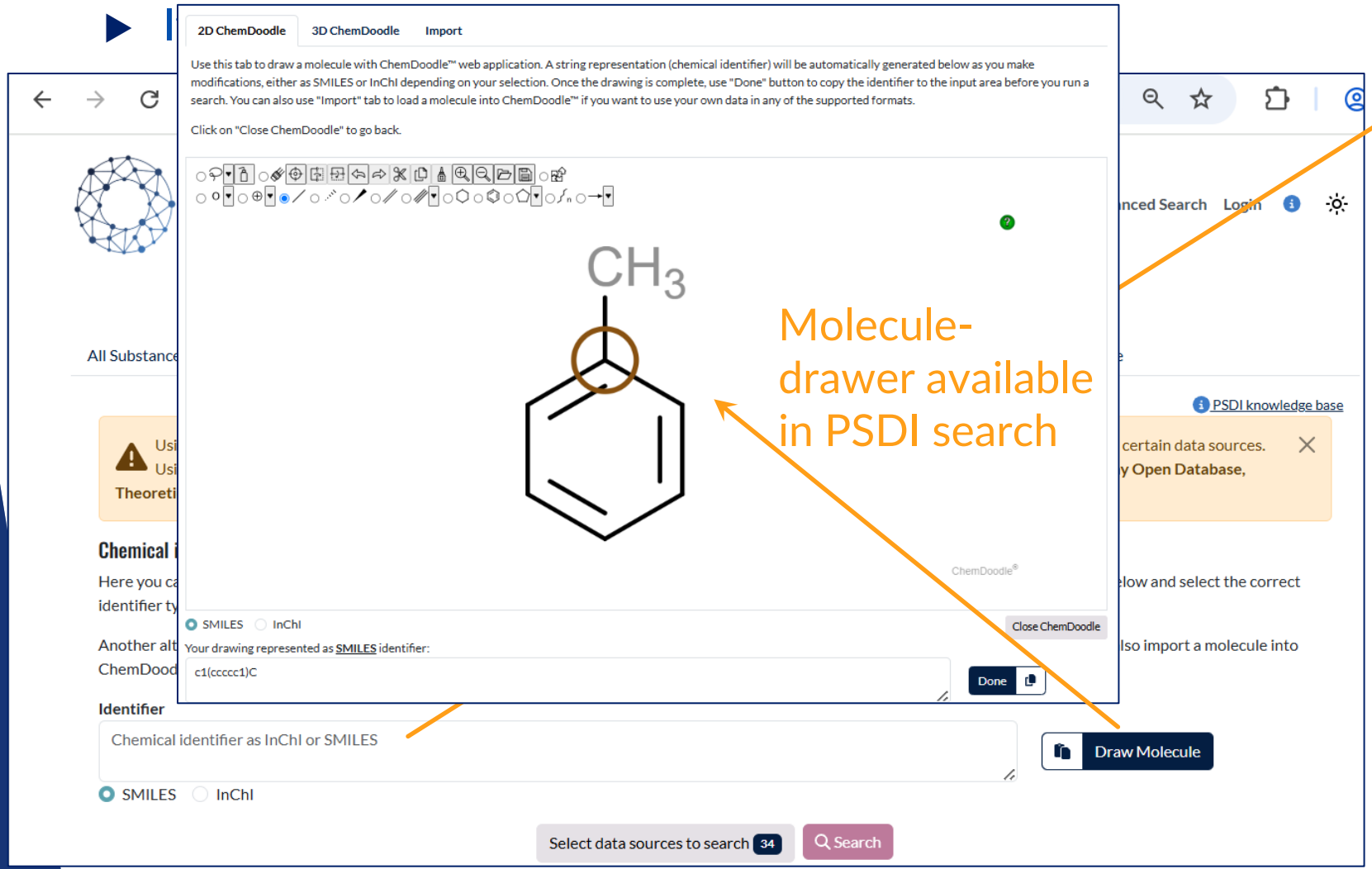
PSDI – find a substance

- If it is a molecule



- searching on custom properties InChI and SMILES) molecular descriptors
- simple text search currently
- There is work on a OPTIMADE cheminformatics namespace for to standardise these

PSDI – find a substance



- searching on custom properties InChI and SMILES) molecular descriptors
- simple text search currently
- There is work on a OPTIMADE cheminformatics namespace for to standardise these