

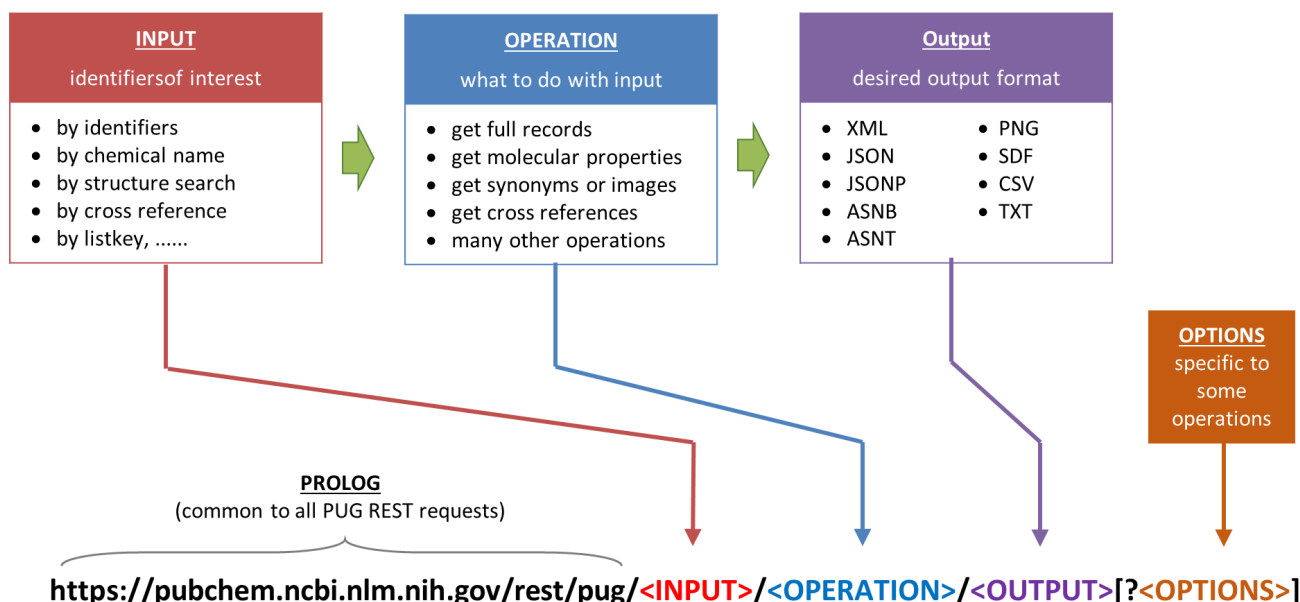
## 1.8: Programmatic Access to the PubChem Database

### Concepts and Syntax of PUG-REST requests

**PUG-REST** is the simplest to use and learn among the existing programmatic access methods to PubChem. Importantly, because information necessary for a **PUG-REST** request can be encoded into a single **Uniform Resource Locator (URL)** that can be written by hand without programming expertise. Conceptually, a web service request from the user to PubChem requires three pieces of information:

- **input:** a list of PubChem identifiers of interest (e.g., CID, AID, SID).
- **operation:** what to do with the input identifiers.
- **output:** the format of the output from the operation.

In PUG-REST, these three pieces of information are encoded into an URL in the following format:



#### [Example]

`https://pubchem.ncbi.nlm.nih.gov/rest/pug/compound/cid/853/record/XML?record_type=3d`

Some tasks require additional pieces of information that do not fit into the three-part PUG-REST URL. They should be provided as a list of ‘&’-separated option name and option value pairs, following the question mark (“?”) appended at the end of the request URL. Some examples are presented in next section, but there are much more things that users can do through PUG-REST. To get more detailed information on PUG-REST, read the following four articles:

- **PUG-SOAP and PUG-REST: Web Services for Programmatic Access to Chemical Information in PubChem**  
Kim *et al.*, *Nucleic Acids Res.* **2015**, 43(W1), W605-W611.  
(<http://dx.doi.org/10.1093/nar/gkv396>).
- **An update on PUG-REST: RESTful interface for programmatic access to PubChem.**  
Kim *et al.*, *Nucleic Acids Res.* **2018**, 46(W1):W563-W570.  
(<https://www.ncbi.nlm.nih.gov/pubmed/29718389>).
- **PUG-REST Help** ([http://pubchem.ncbi.nlm.nih.gov/pug\\_rest/PUG\\_REST.html](http://pubchem.ncbi.nlm.nih.gov/pug_rest/PUG_REST.html))
- **PUG-REST Tutorial** ([http://pubchem.ncbi.nlm.nih.gov/pug\\_rest/PUG\\_REST\\_Tutorial.html](http://pubchem.ncbi.nlm.nih.gov/pug_rest/PUG_REST_Tutorial.html))

#### Example PUG-REST Request for Molecular Properties of a Compound

The request must include the PROLOG, the INPUT, the OPERATION, and the OUPUT. For any request, the prolog will have the format <https://pubchem.ncbi.nlm.nih.gov/rest/pug/>. The INPUT, OPERATION and OUTPUT will change depending on the

context of the information you are requesting.

The **INPUT** is then added. In these three examples, the input is for acetone. the Example inputs can be based on:

- Name **compound/name/acetone/**
- Compound Identifier (CID) **compound/CID/180/**
- InChI Key **compound/inchikey/CSCPPACGZOOCGX-UHFFFAOYSA-N/**

The **OPERATION** is then added. In this case we will get the molecular weight, molecular formula, and its SMILES line notation string:**property/MolecularWeight,MolecularFormula,CanonicalSMILES/**

The **OUTPUT** can be obtained as text or comma separated values or eXtensible Markup Language data.

- Text= **TEXT NOTE:** *This output type is limited to a single property value*
- comma separated values= **CSV**
- eXtensible Markup Language= **XML**

Putting these all together results in the following example requests:

1. <https://pubchem.ncbi.nlm.nih.gov/rest/pug/compound/name/acetone/property/MolecularWeight,MolecularFormula,CanonicalSMILES/XML>
2. <https://pubchem.ncbi.nlm.nih.gov/rest/pug/compound/CID/180/property/MolecularWeight/TXT>
3. <https://pubchem.ncbi.nlm.nih.gov/rest/pug/compound/inchikey/CSCPPACGZOOCGX-UHFFFAOYSA-N/property/MolecularWeight,MolecularFormula,CanonicalSMILES/CSV>

### Try it yourself!

Write a PUG-REST URL Request that returns an XML file for morphine that contains values for its compound identifier, IUPAC name, molecular formula, and hydrogen bond acceptor sites in the molecule. (*Hint: look at the output for example 1 above.*)

---

1.8: Programmatic Access to the PubChem Database is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.