

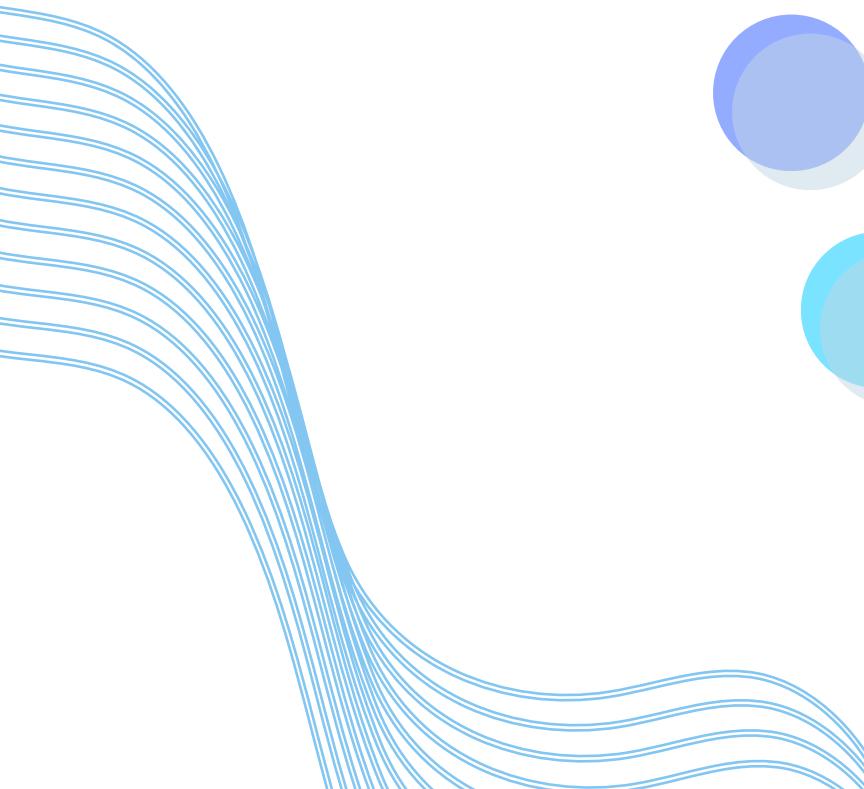
RNA LIBRARY



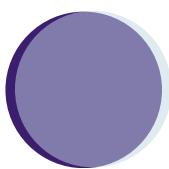
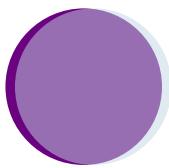
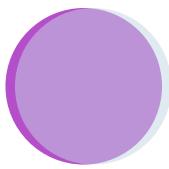
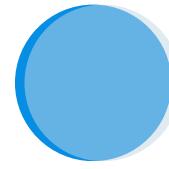
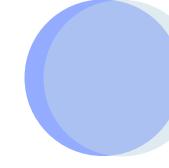
Presented by: Rayane Adam and Joelle Assy
Object-Oriented Programming 2, 2024-2025

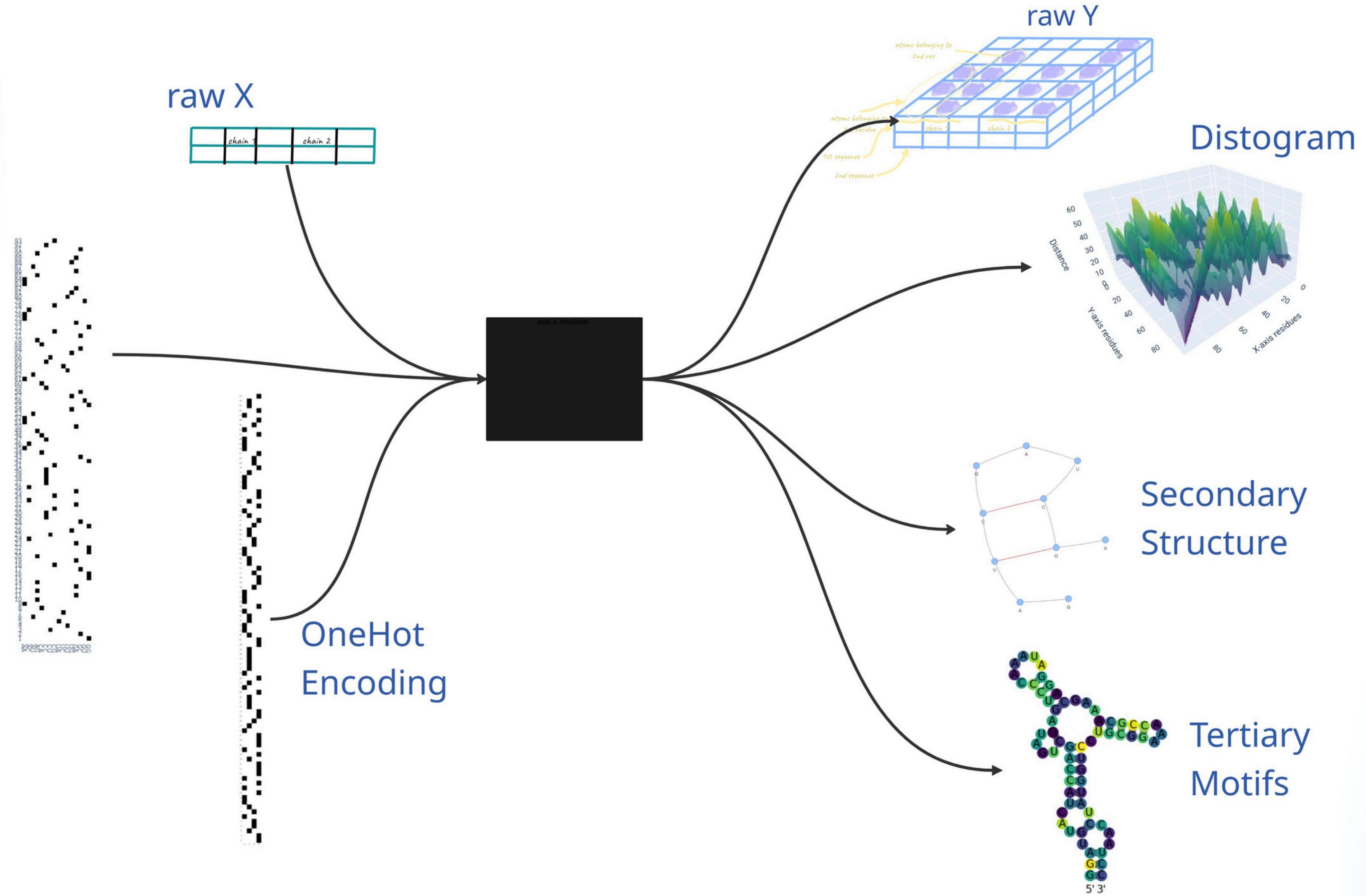
Table of Content

- Overview of Library Functionalities
- Demo
- UMLs
- Modules and Design Patterns
- Transformations and Visualization
- Usage/Installation/Documentation/Extensions

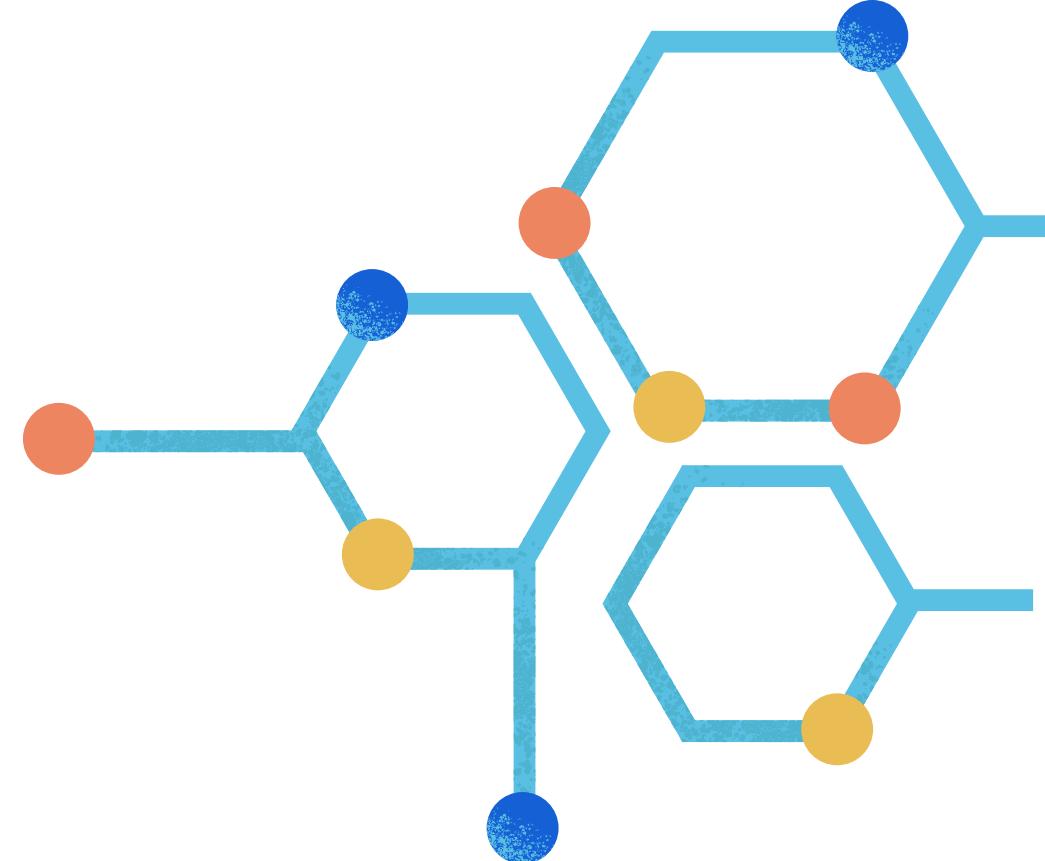
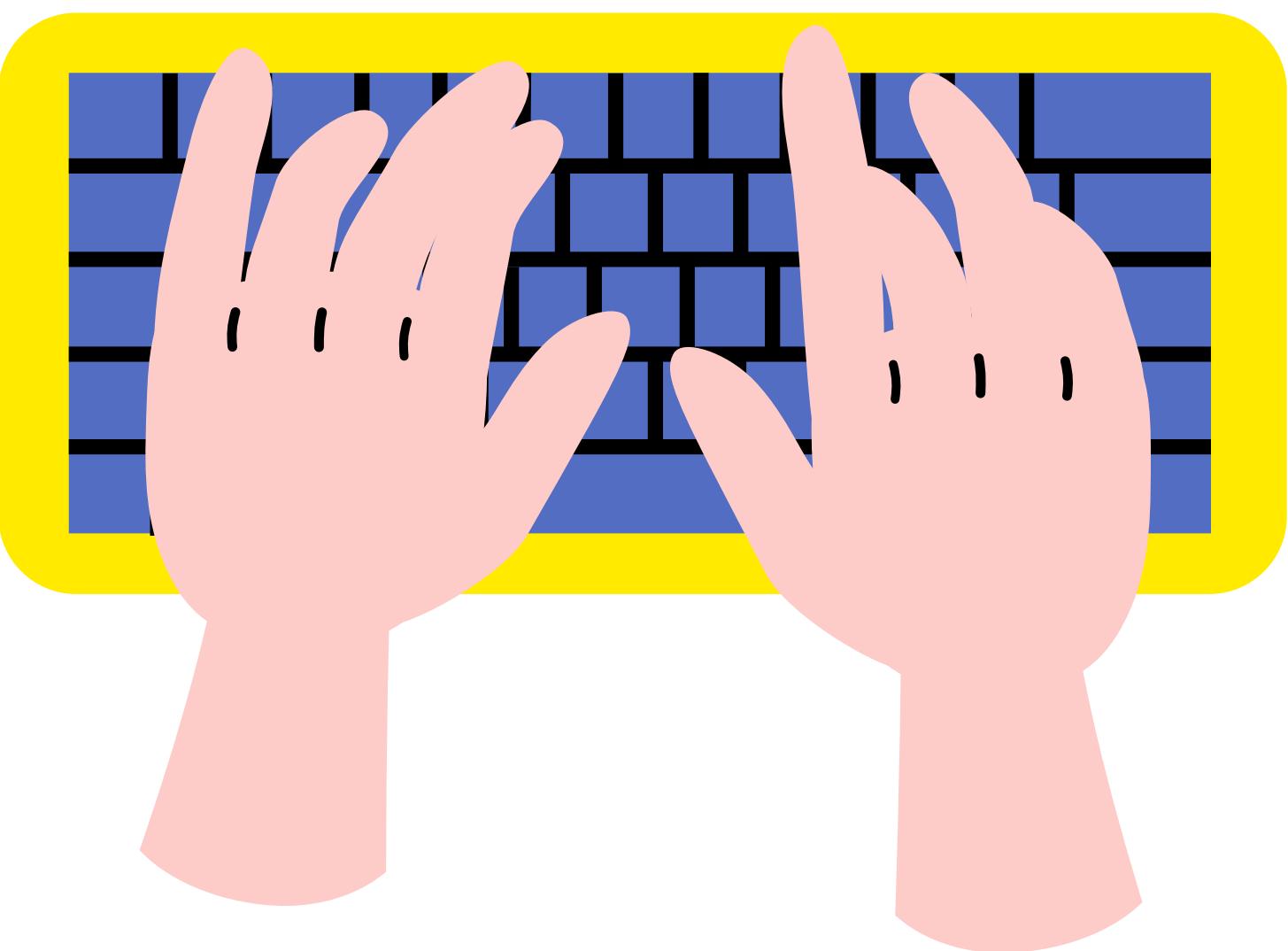
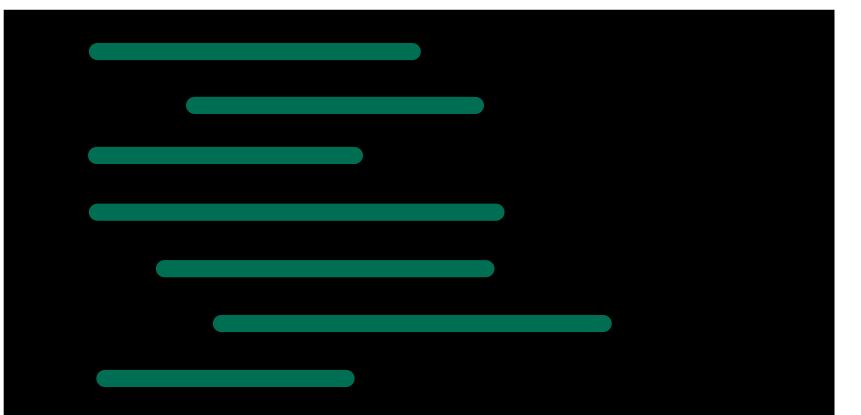


LIBRARY FUNCTIONALITIES

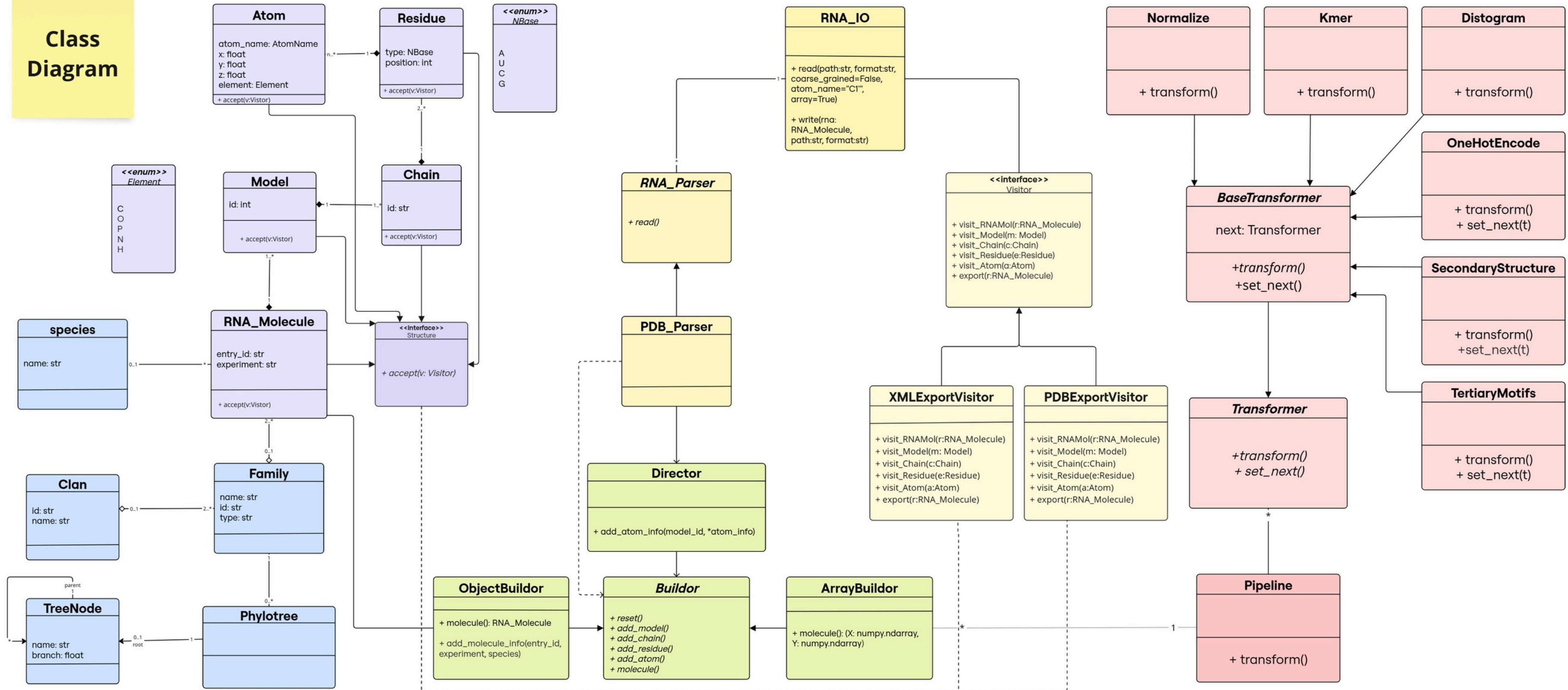
-  Creating and manipulating RNA molecules objects.
-  Reading PDB file and creating an RNA molecule object or a numpy array of coordinates.
-  Writing PDB and XML files from an RNA molecule object.
-  Reading multiple PDB files into a single numpy array of sequences, and another numpy array of coordinates.
-  Creating a Pipeline of Transformations on the sequences and coordinates arrays.
-  Visualizations



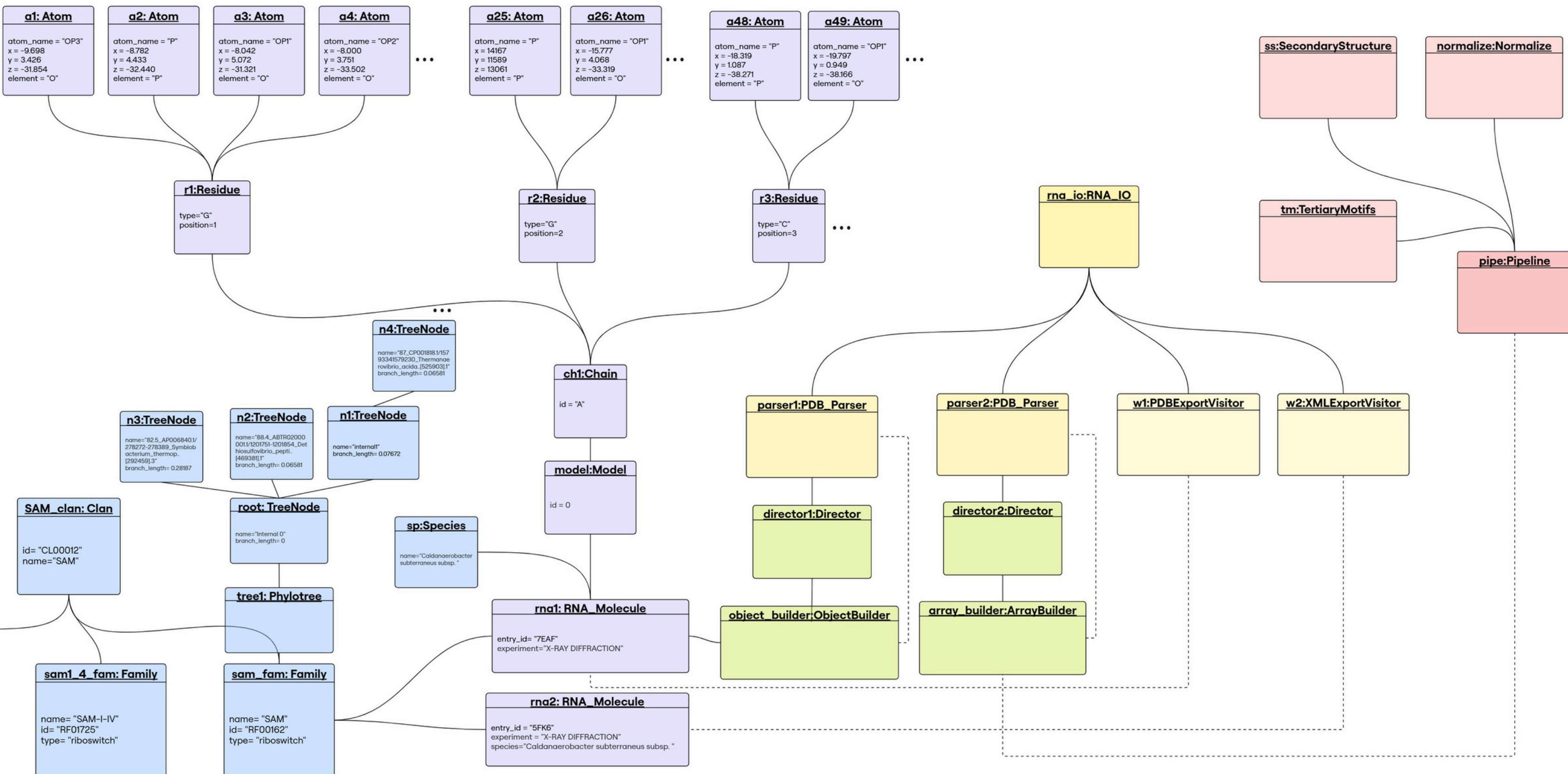
Demo



Class Diagram



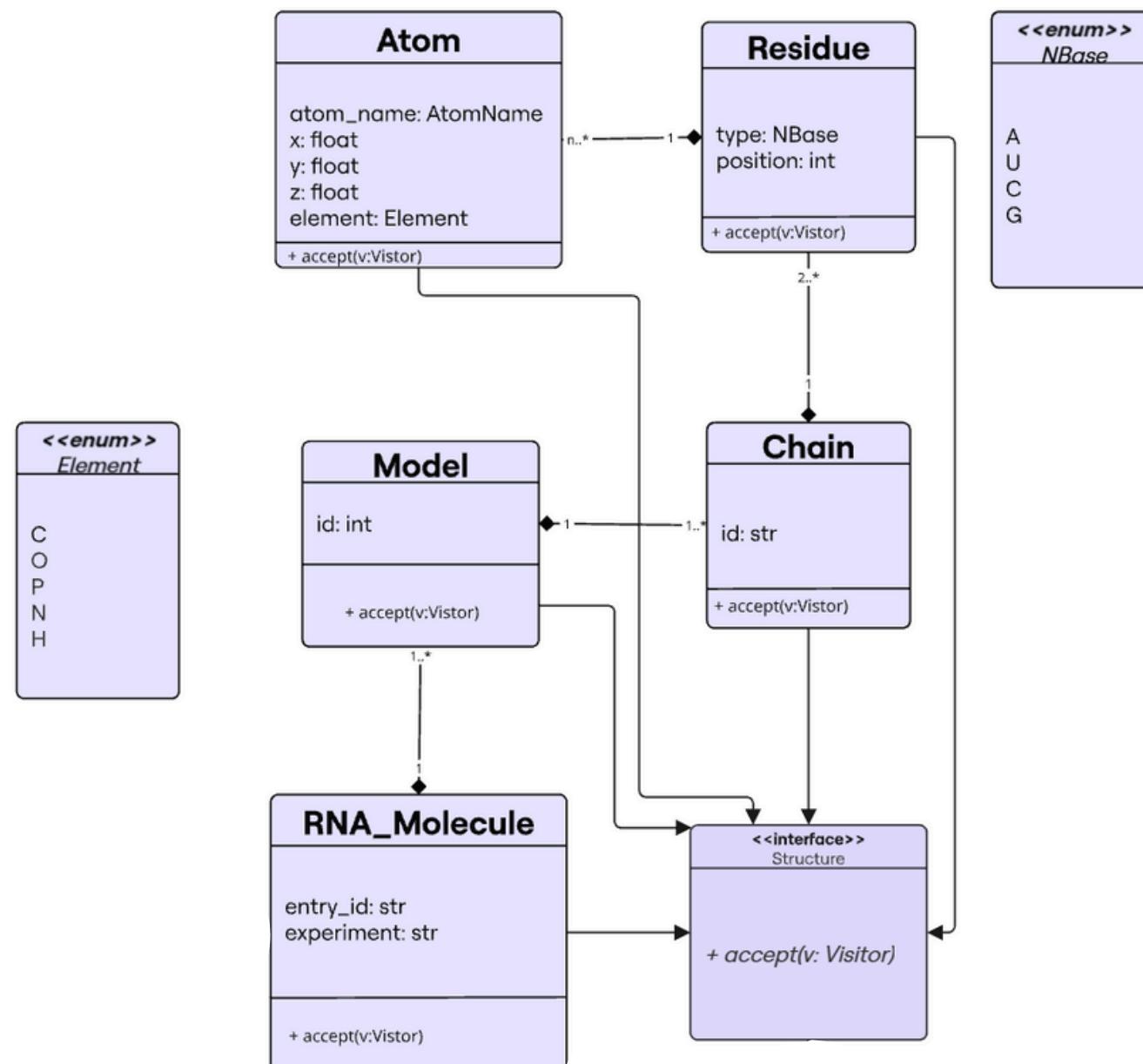
Object Diagram



Structure Module

Common Implementation

- Getters and setters for data integrity and type validation.
- `__repr__` for clear object output.
- Parents store children in dictionaries for easy access.
- Children reference parents for bidirectional navigation.
- Parents manage adding, removing, and getting children.
- Objects can be initialized with existing children.

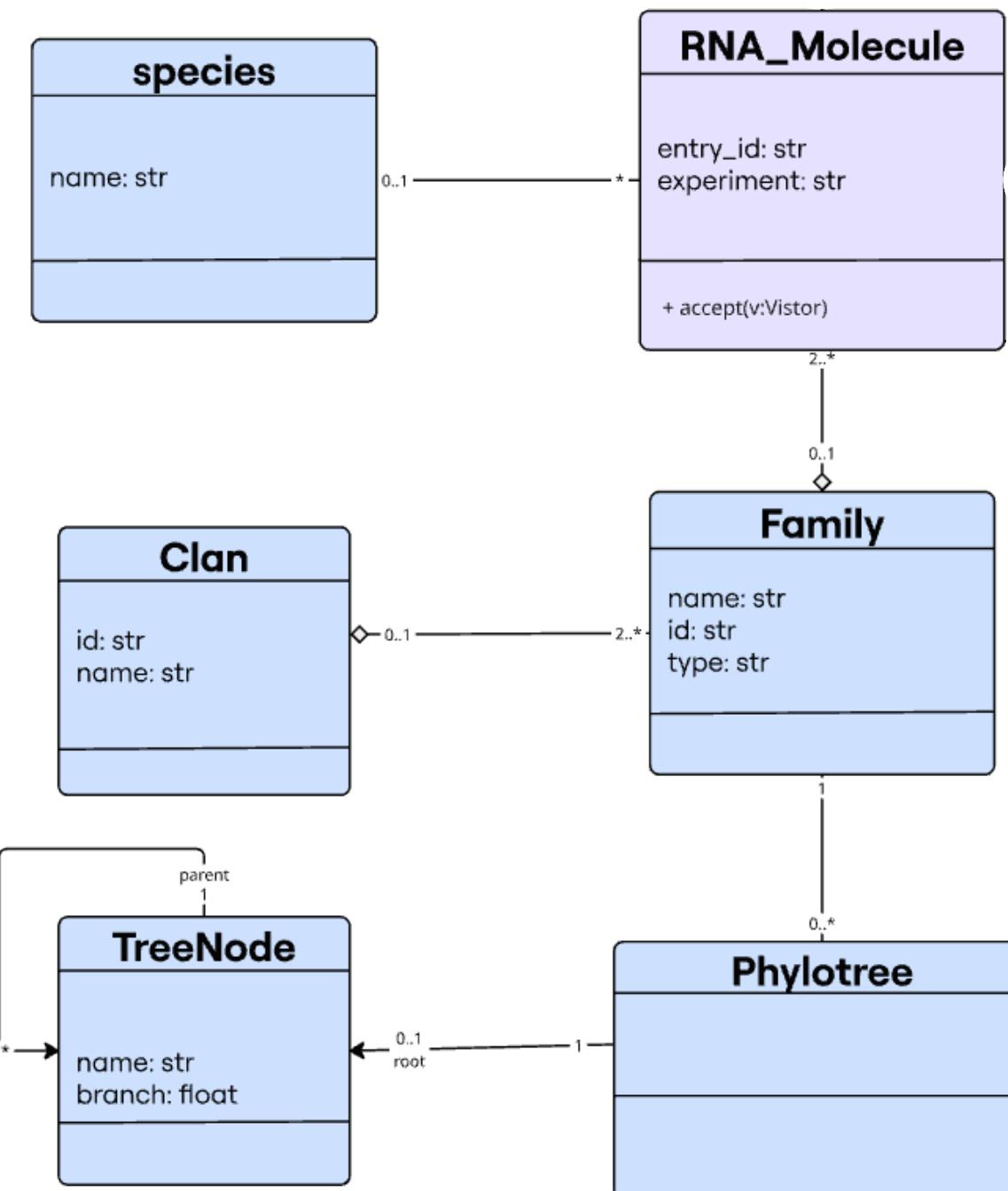


Families Module

Families of RNA

Describe different entities associated with RNA evolutionary relationships between species

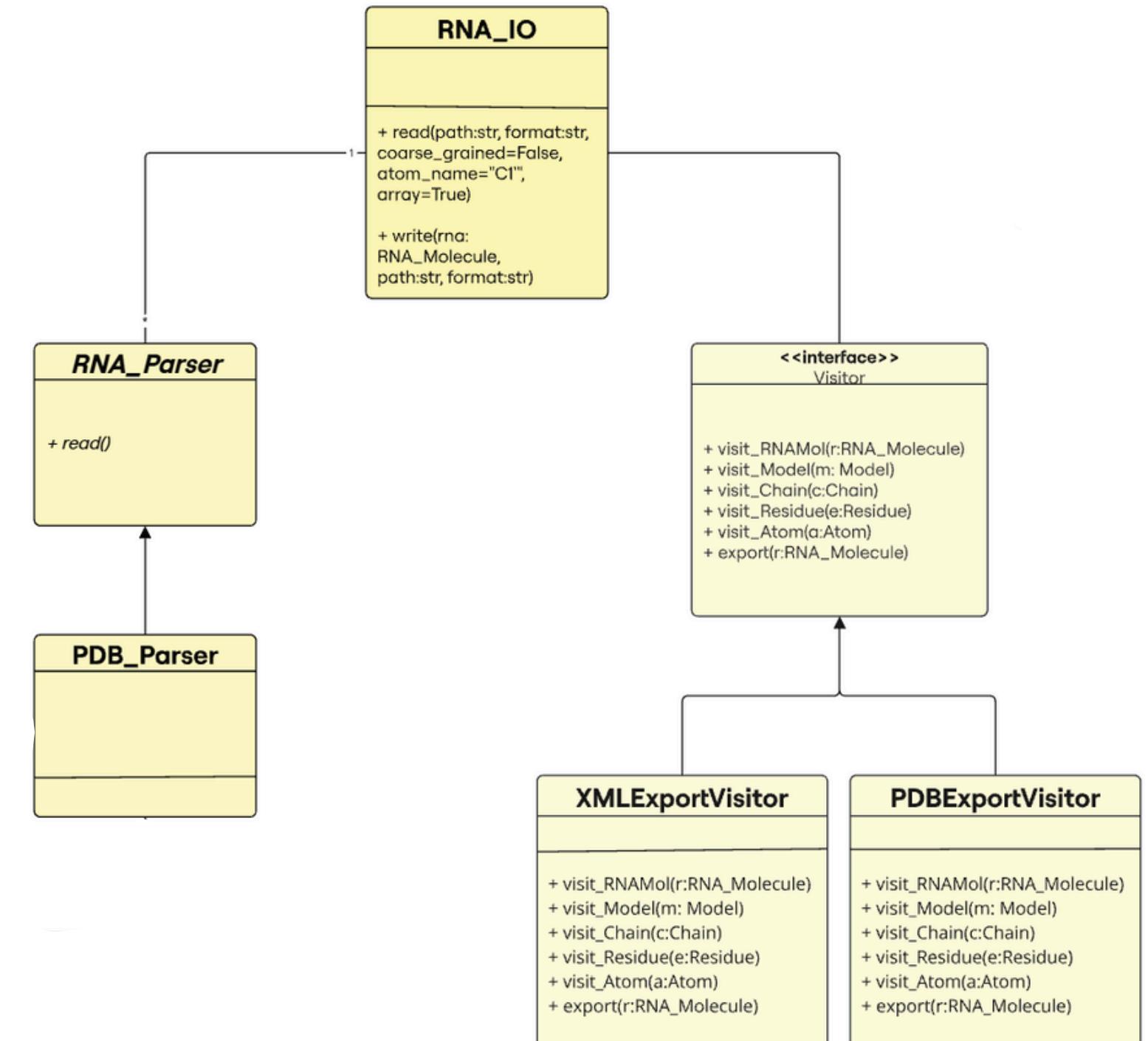
- Supports access to real entities from rfam database
- Dunder magic methods and private methods added for flexibility with supported encapsulation
- Class attributes added to ensure one instance of one entity, like in database design



IO Module

PDB Parser

- Uses the Builder pattern (Director & Builder classes) for structured data construction.
- Extracts PDB metadata (ID, experiment type, species) via `extract_molecule_info()`.
- Extracts atom info (atom attributes, residue attributes, chain attribute via `extract_atom_info()`).
- Supports multi-model structures, assigning atoms to their respective models.



IO Module

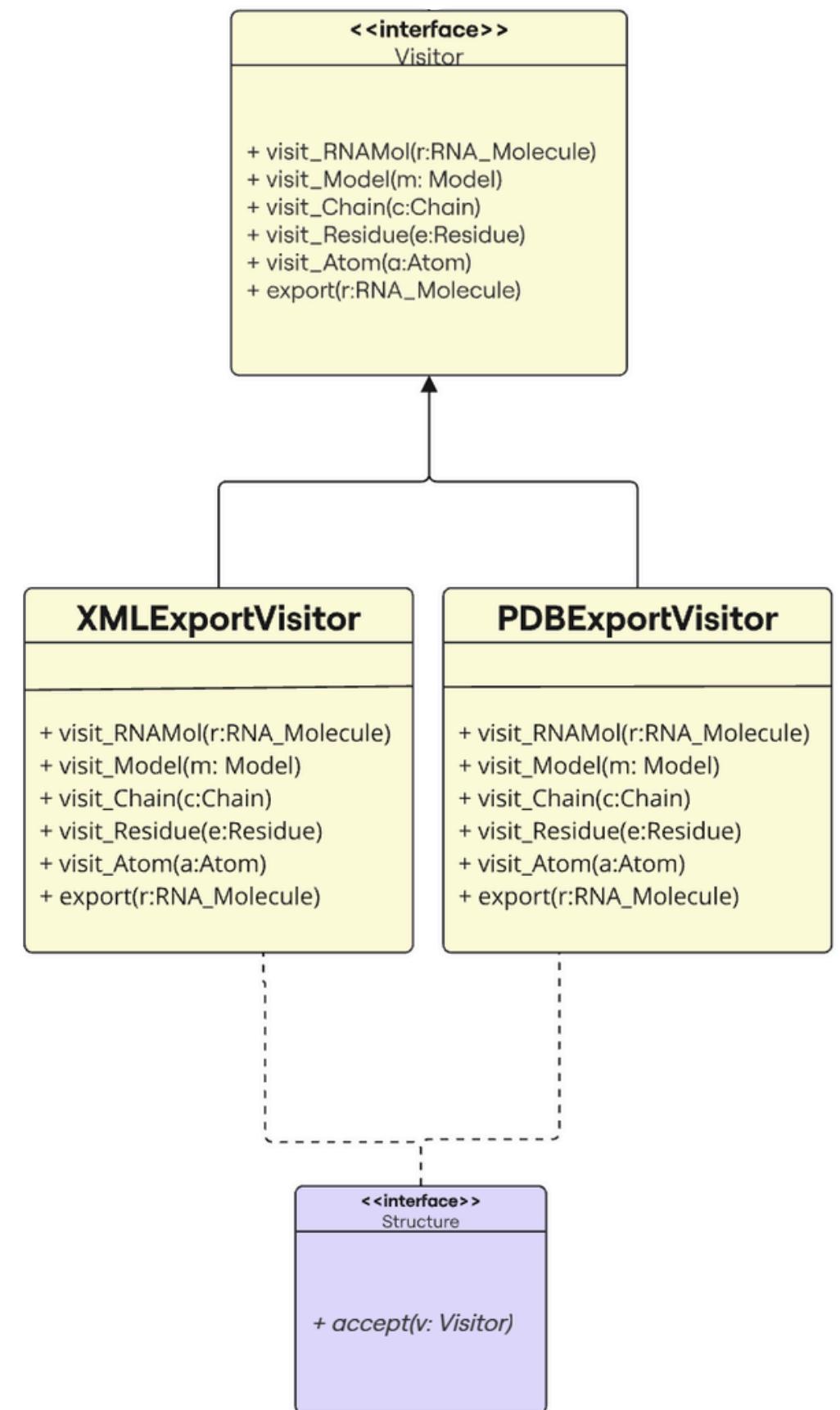
Visitor Design Pattern

Advantages:

- Separates the logic of traversing the molecular structure from the operations applied to it
- Supports adding new output formats (e.g., CSV, JSON)

Disadvantages:

- higher complexity



Processing Module

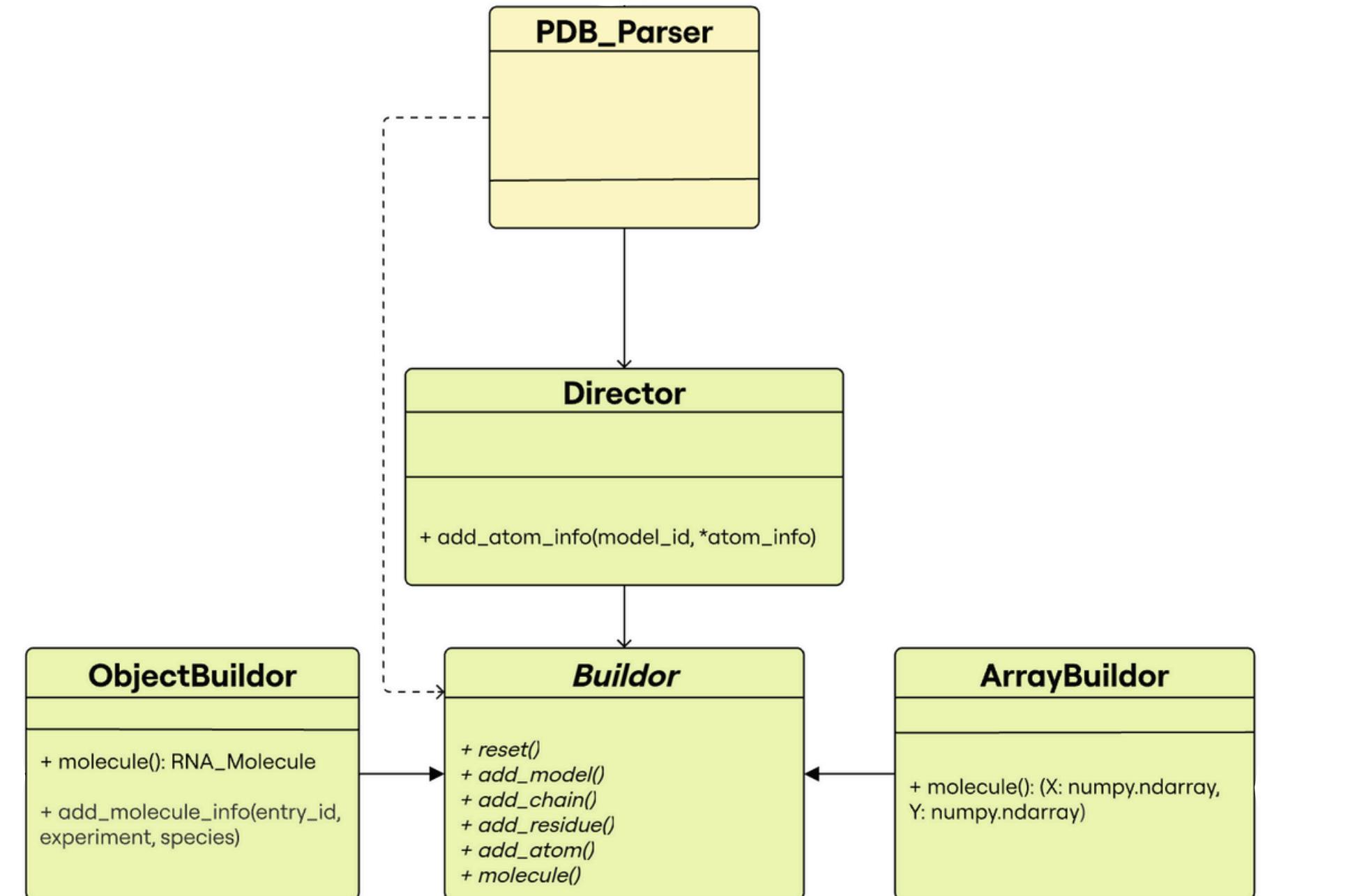
Builder Design Pattern

Disadvantages:

Adds complexity

Advantages:

- Dedicated builder class
- Direct construction
- Common construction recipe
- Smaller steps in construction

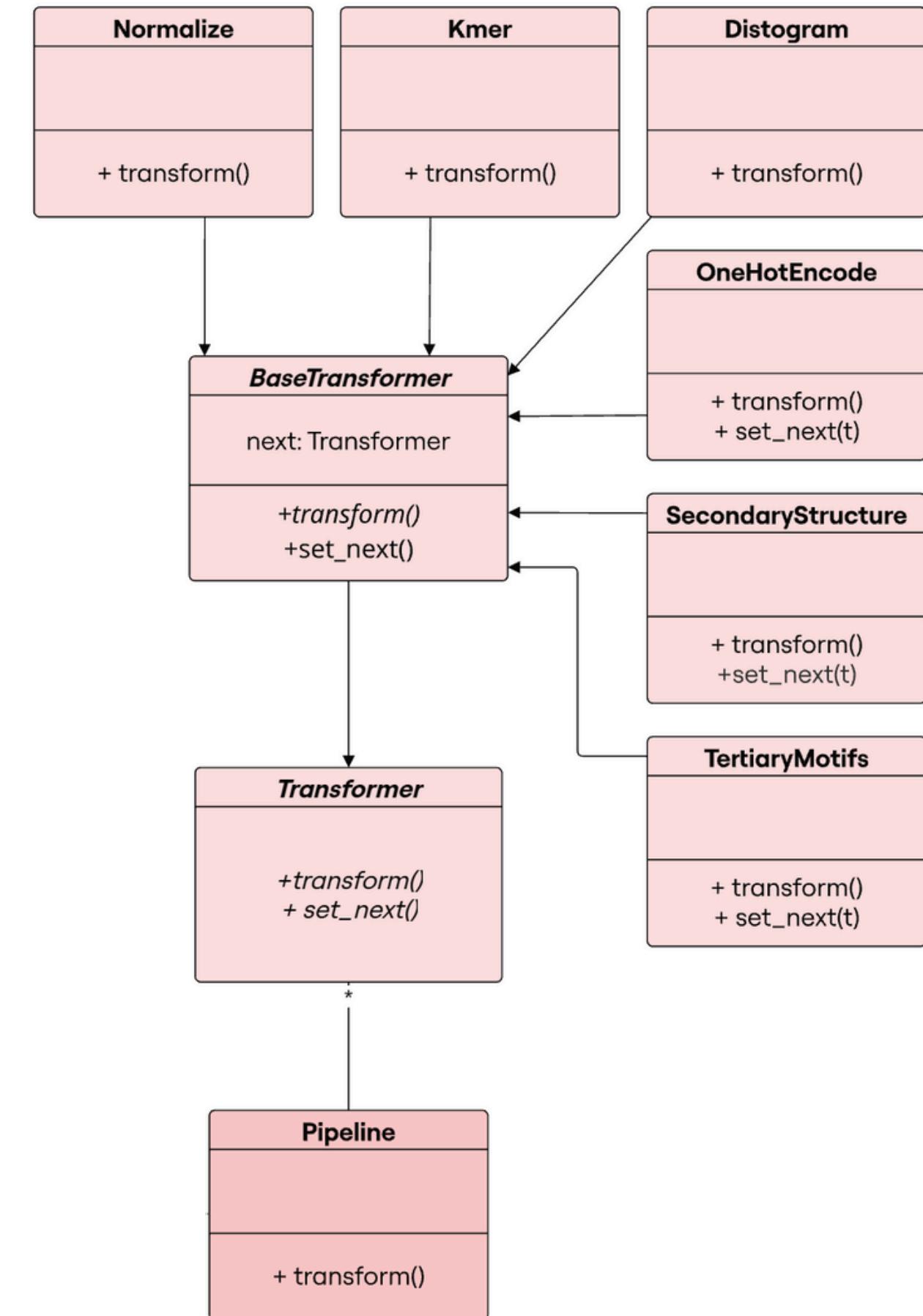


np_sequence: (models, max_residues)

np_array: (models, max_residues, max_atoms, 3)

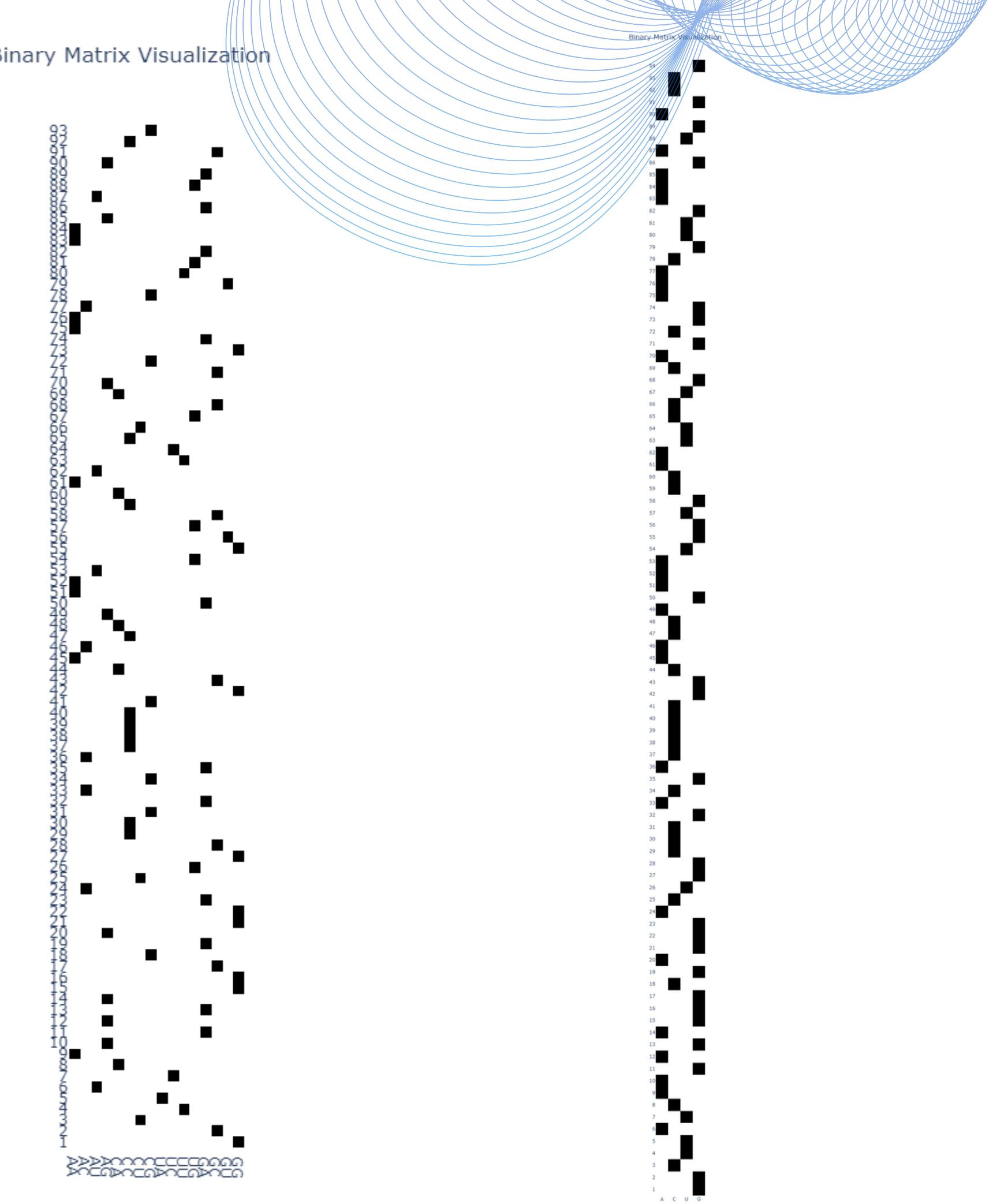
Transformations Module

Chain of Responsibility
Design Pattern



ONEHOTENCODE

- Can be output of
 - raw sequence
 - kmerized sequence



3D vis distogram (94x94) landscape

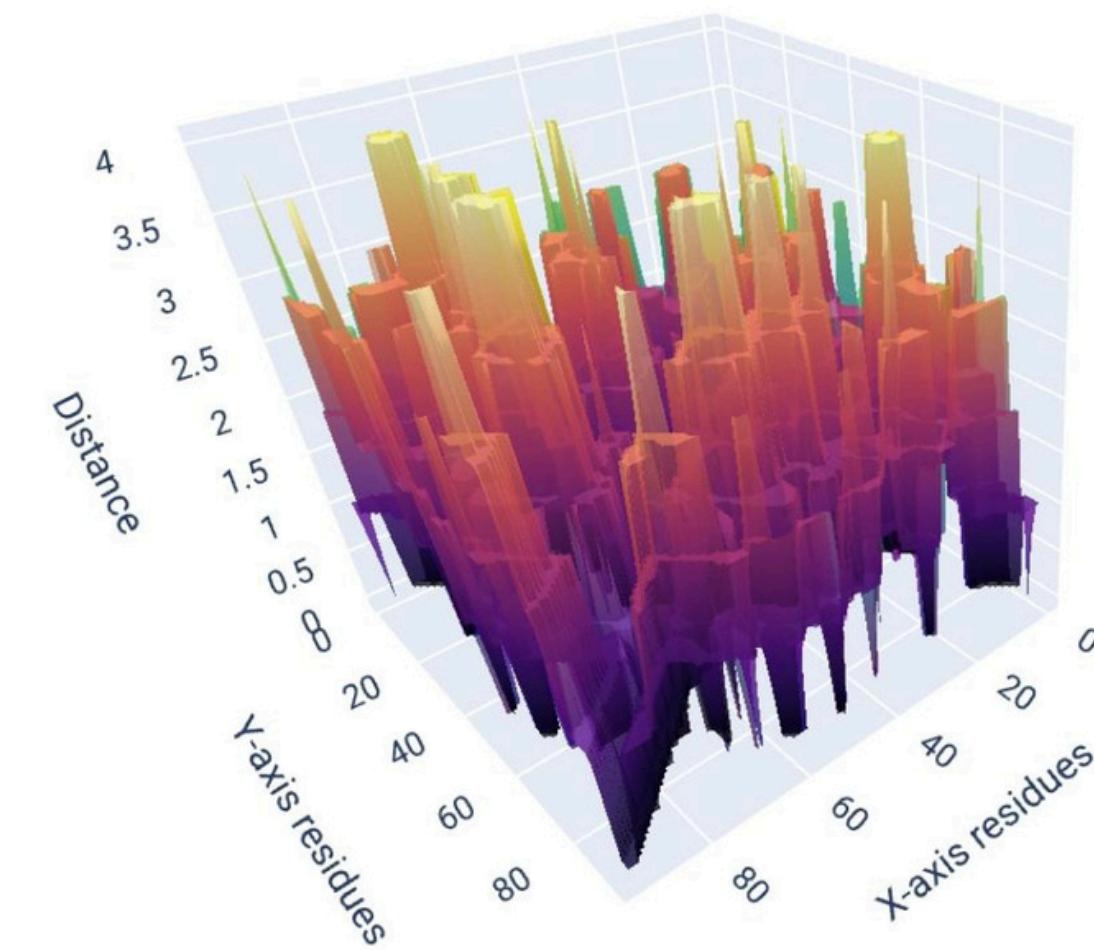
DISTOGRAM

- Distance matrix
- Can be of different dimensions, depending on:
 - number of atoms
 - Number of buckets

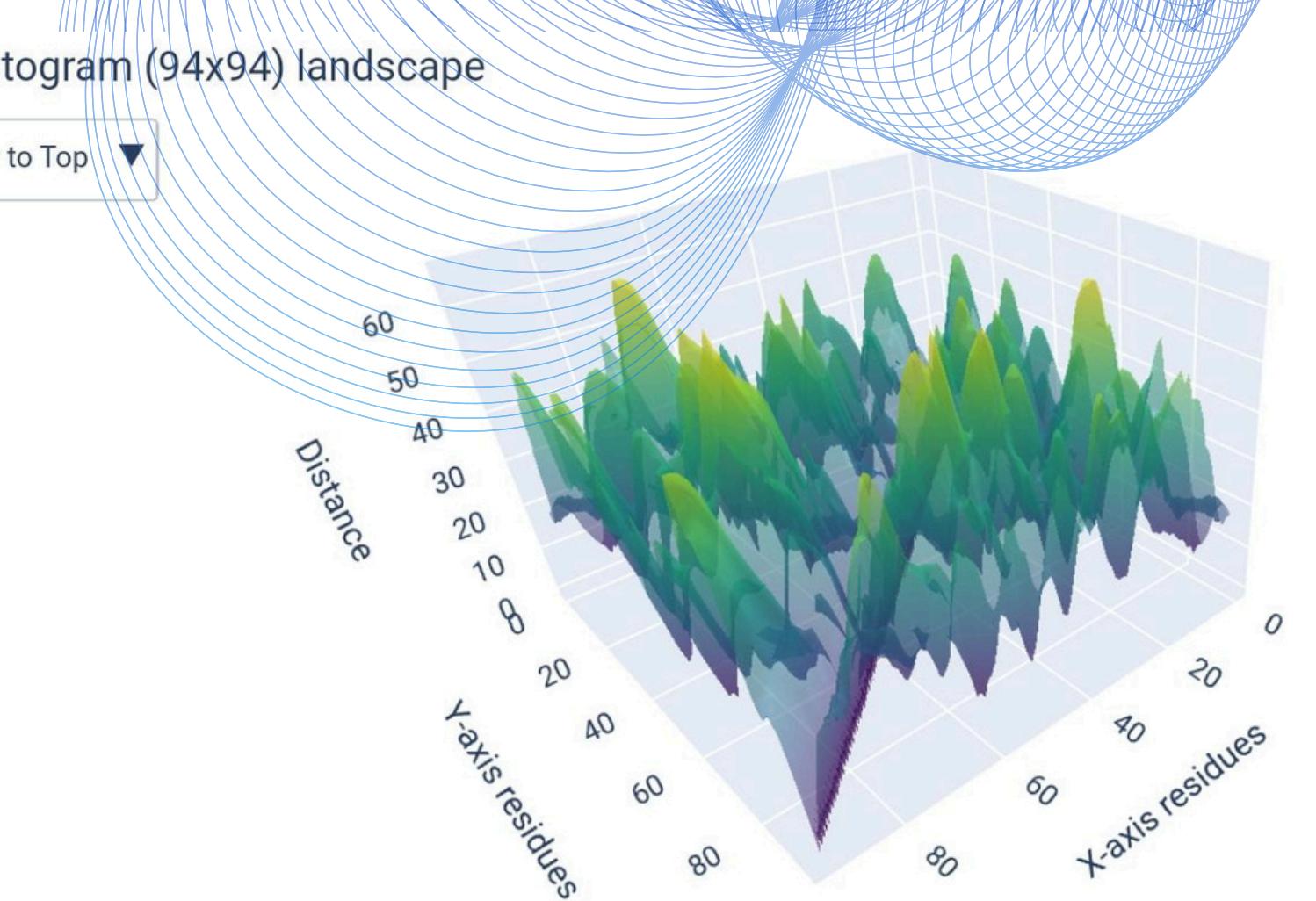
Output from coordinates

3D vis distogram (94, 94, 3, 5) landscape discretized

Bring atom 2 in list to Top ▼

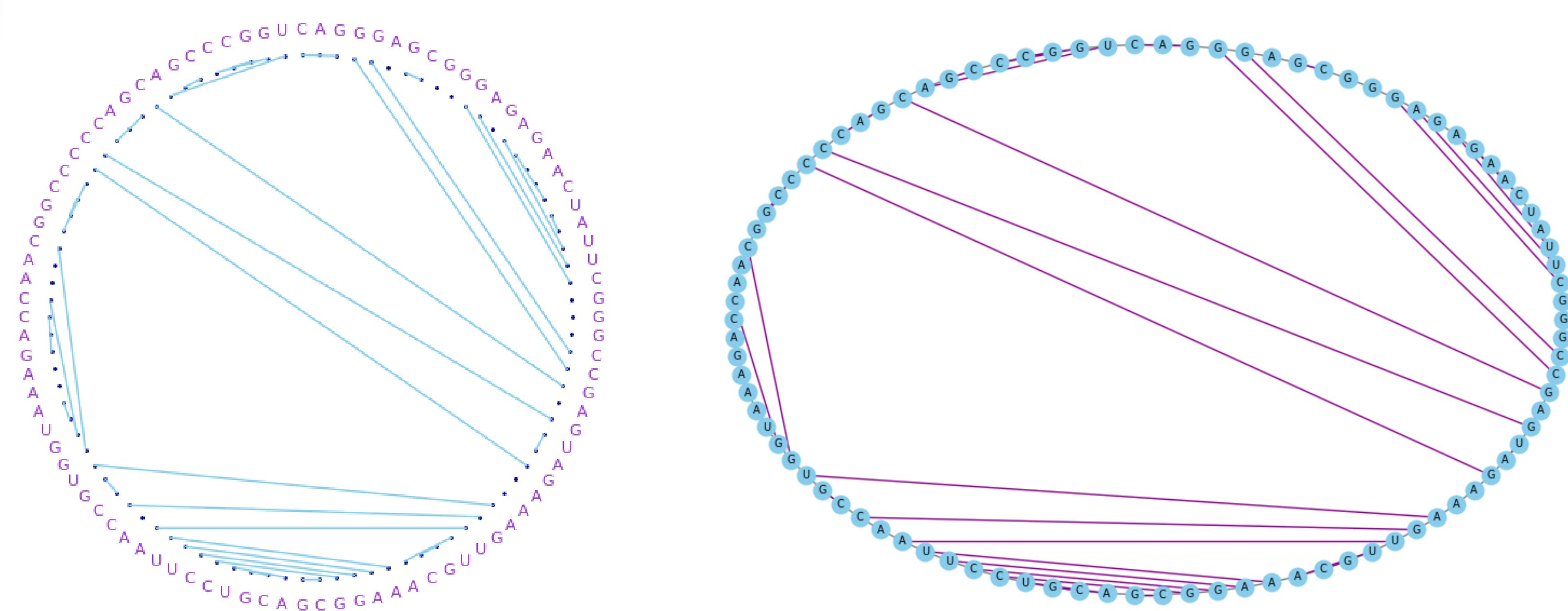


atom 1 in list layer to Top ▼



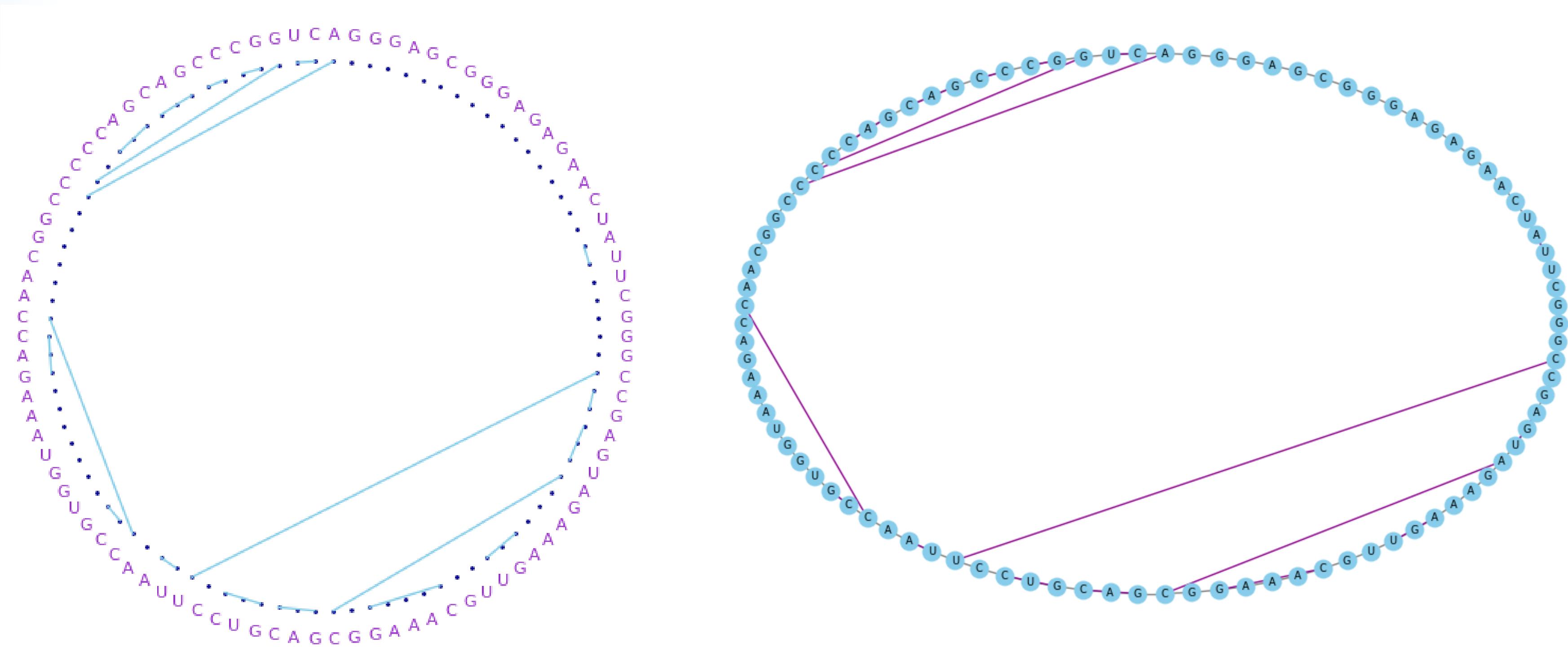
SECONDARY STRUCTURE

Nussinov Algorithm



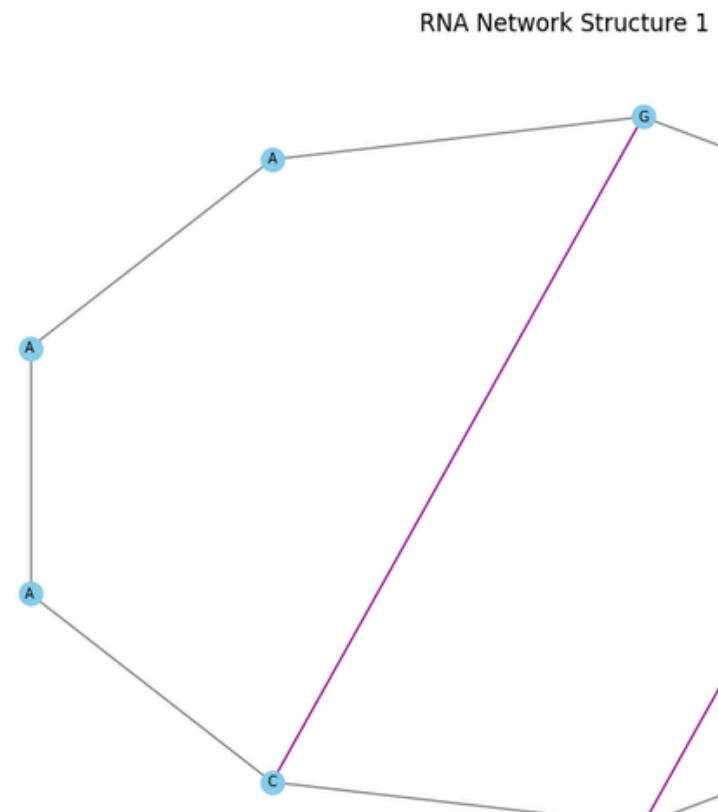
SECONDARY STRUCTURE

Watson Crick Distances

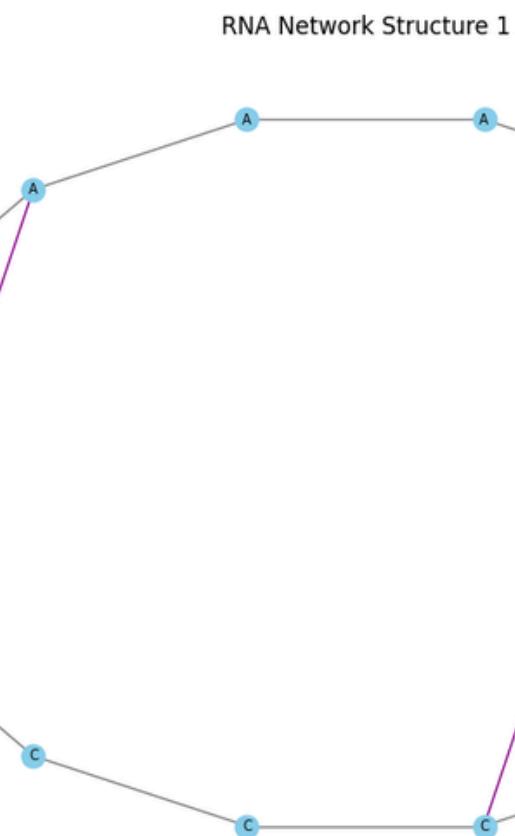


TERTIARY MOTIFS

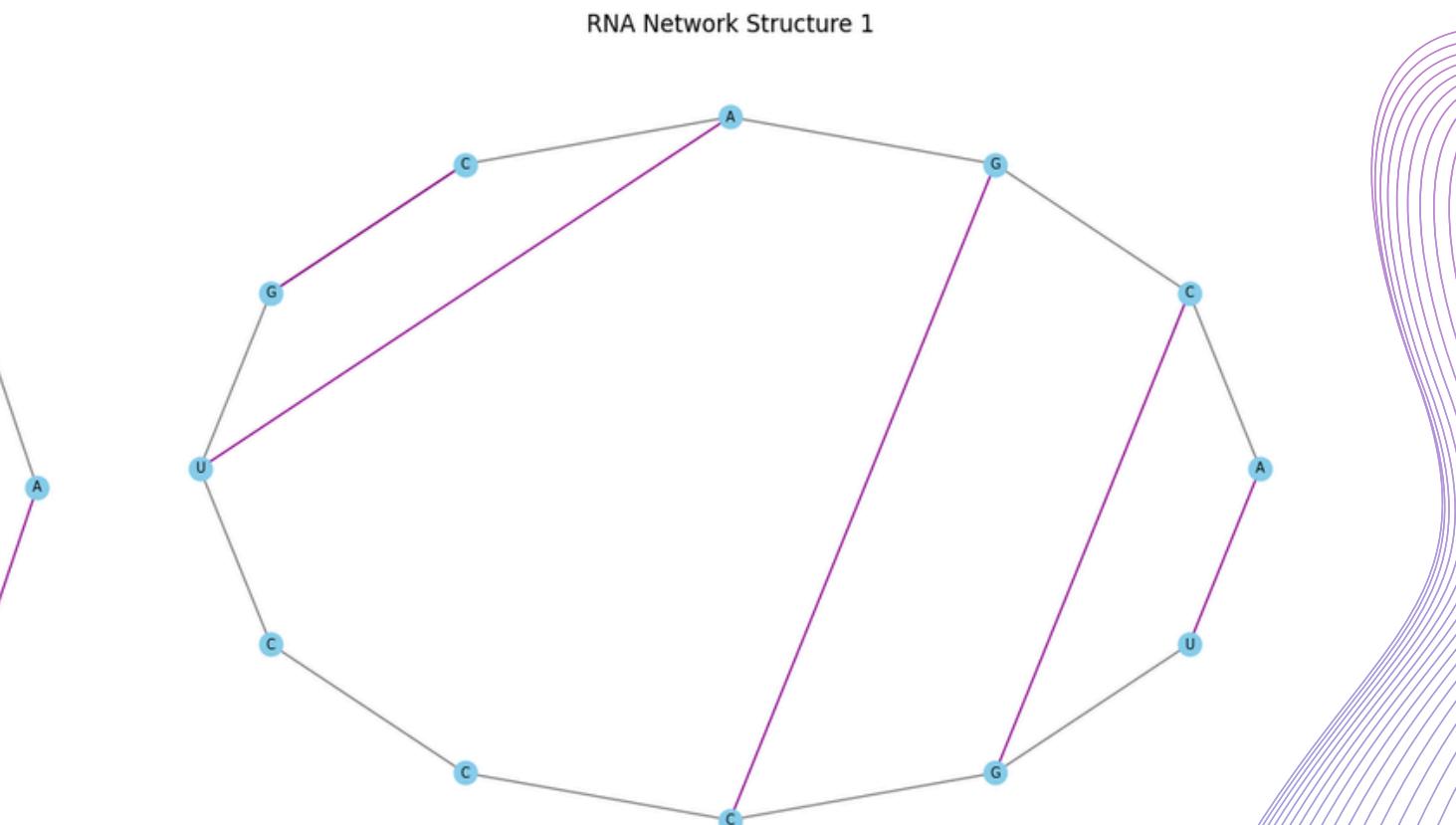
Hairpins



Internal_loops



Bulges



Extensions

- Add more algorithms related to 2ary and 3ary structure prediction
- provide multi-classification machine learning support or ensemble learning by providing RNA specific loss functions
- Train and test on real data

rnalib

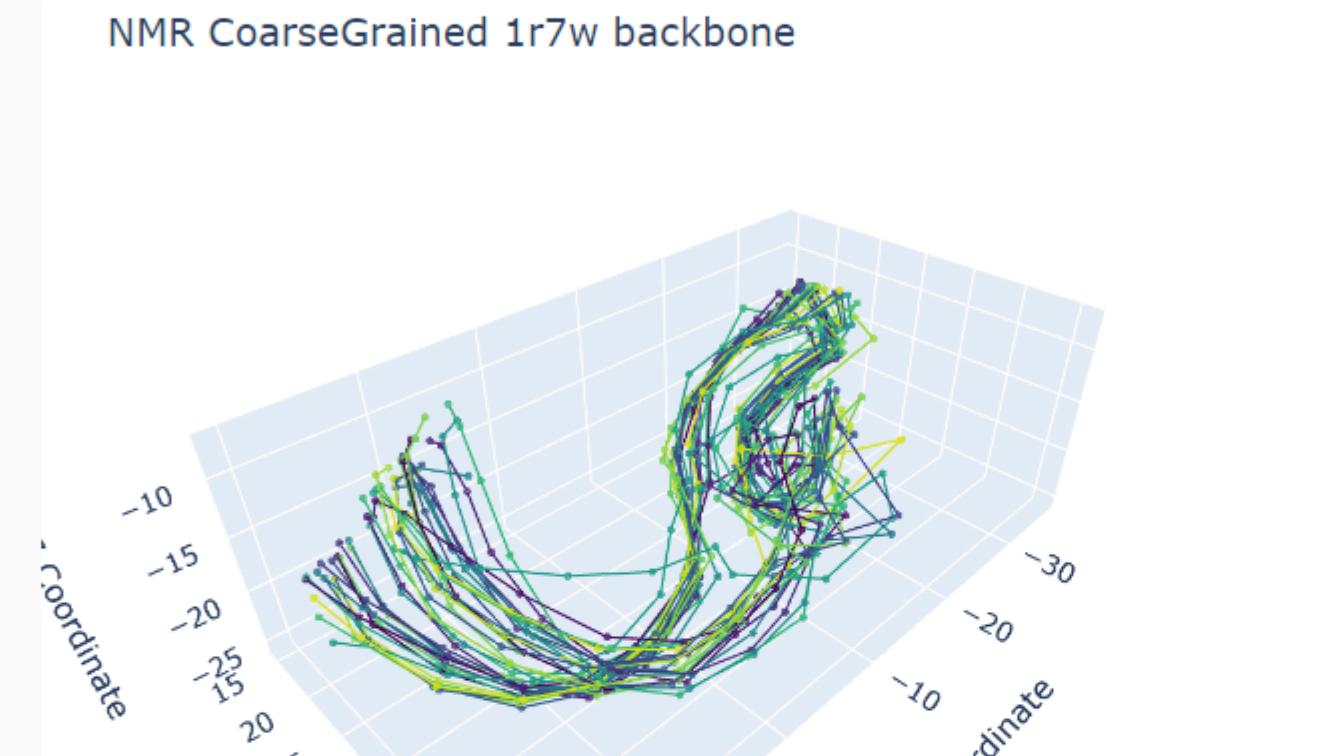
Search docs

CONTENTS:

- rnalib

rnalib package

NMR CoarseGrained 1r7w backbone



Indices and tables

- Index
- Module Index
- Search Page

This library can be installed directly from github through pip:

```
pip install git+git://github.com/rna-oop/2425-m1genomhe-group-6.git
```

Online demo available as soon as it becomes public: [Open in Colab](#)

Documentation of the various functions and classes can be found in a [Sphinx generated read-the-docs-like webpage](#) or in the docstrings of the code. Not deployed on github pages yet because the repo is private.

© Copyright 2025, m1genomhe2425-rna-oop-group6.

Built with Sphinx using a theme provided by Read the Docs.

Thank You