

Flowcept Agent Chat



Hi, there! I'm a **Workflow Provenance Specialist**.

I am tracking workflow executions and I can:

- 🔍 Analyze running workflows
- 📊 Plot graphs
- 🗨️ Answer general questions about provenance data

How can I help you today?

💡 Tip: Ask about workflow metrics, generate plots, or summarize data.



from this molecule, which bond has the highest dissociation free energy?



Here's the code:

```
result = df[df['activity_id'] == 'run_individual_bde'][['used.frag_label',
```

	used.frag_label	generated.bd_free_energy
0	C-H_2	92.3911



Summary:

1. The DataFrame contains two columns: `used.frag_label` and `generated.bd_free_energy`. The `used.frag_label` column likely represents the label of a fragment or bond type within a molecule, and `generated.bd_free_energy` represents the dissociation free energy of that bond.
2. The `Query_Code` is designed to filter the DataFrame for entries where the activity ID is 'run_individual_bde', suggesting it's looking for data related to bond dissociation energies. It then selects the columns for bond labels and their corresponding dissociation free energies, sorts them in descending order based on the energy, and retrieves the top result. This means the result will show the bond with the highest dissociation free energy.
3. The bond with the highest dissociation free energy in the molecule is the C-H bond, with a dissociation free energy of 92.391084 kcal/mol.

Send a message





What functional was used for the calculations?



Here's the code:

```
result = df[[col for col in df.columns if col.startswith('generated.function
```

	generated.functional
7	B3LYP
8	B3LYP
9	B3LYP
10	B3LYP
11	B3LYP
12	B3LYP
13	B3LYP
14	B3LYP
15	B3LYP
16	B3LYP

 Summary:

The functional used for the calculations is B3LYP.

Send a message





In this molecule, what basis set was used for the calculation?



Here's the code:

```
result = df[df['activity_id'] == 'write_nwchem_input'][['generated.basis_set',
```

	generated.basis_set
0	6-31G*
1	6-31G*
2	6-31G*
3	6-31G*
4	6-31G*
5	6-31G*
6	6-31G*
7	6-31G*
8	6-31G*
9	6-31G*

 Summary:

The basis set used for the calculation in the molecule is 6-31G*.

Send a message





In this molecule, what is the lowest energy bond enthalpy?



Here's the code:

```
result = df[df['activity_id'] == 'run_individual_bde'][['generated.bd_entha'
```

	generated.bd_enthalpy
0	83.5652

 Summary:

1. The DataFrame contains a single column labeled 'generated.bd_enthalpy' which represents bond enthalpy values in some unit, likely kilojoules per mole (kJ/mol). The sampled row shows a bond enthalpy value of 83.565186.
2. The query code is designed to filter the DataFrame for entries related to the activity 'run_individual_bde' and then extract the 'generated.bd_enthalpy' column. It subsequently finds the minimum value of this column, which represents the lowest bond enthalpy in the dataset.
3. The lowest energy bond enthalpy in the molecule is 83.565186 kJ/mol.

Send a message





In this molecule, what is the number of atoms in the molecule?



Here's the code:

```
result = df[df['activity_id'] == 'mol_to_xyz'][['generated.num_atoms']]
```

	generated.num_atoms
0	9
1	4
2	5
3	7
4	2
5	8
6	1
7	8
8	1
9	8

 Summary:

1. The DataFrame contains a single column labeled 'generated.num_atoms' which lists the number of atoms in different molecules. The values in this column are all numerical and represent the count of atoms in each molecule.
2. The `Query_Code` specifically filters the DataFrame for entries where the 'activity_id' equals 'mol_to_xyz', and then selects the column 'generated.num_atoms'. This suggests that the data represents the number of atoms in molecules that have been converted to XYZ format (a common format for representing molecular structures in 3D space).
3. The number of atoms in the molecule varies as shown in the DataFrame, with values such as 9, 4, 8 atoms, etc. Each row represents a different molecule.

The number of atoms in the molecule varies; examples include 9, 4, and 8 atoms.

Send a message





In this molecule, what is the multiplicity and charge of the parent molecule?



Here's the code:

```
result = df[['generated.species.parent.multiplicity', 'generated.species.pa
```

	generated.species.parent.multiplicity	generated.species.parent.charge
0	1	0



Summary:

The parent molecule has a multiplicity of 1 and a charge of 0.

Send a message





Please plot a bar graph displaying the bond dissociation enthalpy for for each bond label?



Here's the code:

```
result = df[['generated.bond_id', 'generated.bd_enthalpy']].dropna().groupby
```

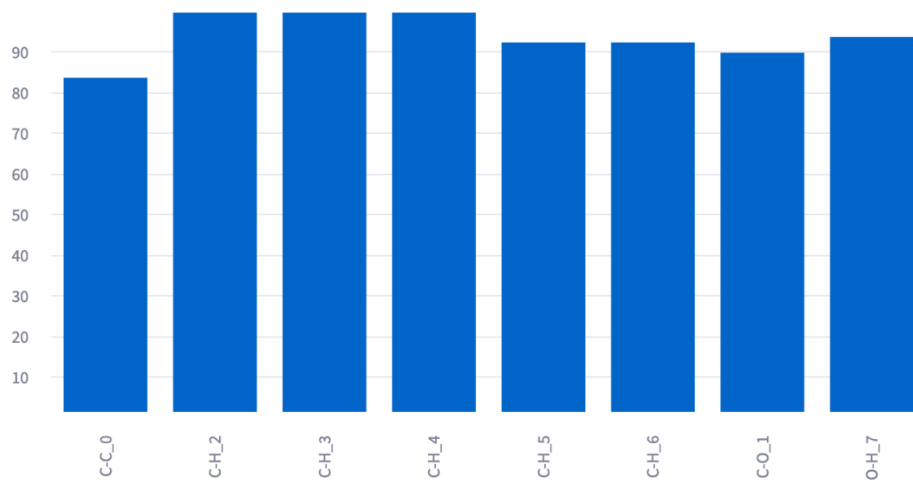
	bond_id	average_bd_enthalpy
0	C-C_0	83.5652
1	C-H_2	100.2277
2	C-H_3	100.2277
3	C-H_4	100.2277
4	C-H_5	92.2554
5	C-H_6	92.2554
6	C-O_1	89.7186
7	O-H_7	93.6152

Here's the plot code:

```
st.bar_chart(result.set_index('bond_id'))
```



Here's the plot:



Send a message



Flowcept Agent Chat


 Hi, there! I'm a **Workflow Provenance Specialist**.

I am tracking workflow executions and I can:

- 🔍 Analyze running workflows
- 📊 Plot graphs
- 🗨️ Answer general questions about provenance data

How can I help you today?

💡 Tip: Ask about workflow metrics, generate plots, or summarize data.

 In the current workflow, what is the average bond dissociation enthalpy for the bond labels that contain 'C-H'?

 📊 Here's the code:

```
result = df[df['activity_id'] == 'run_individual_bde'][df['generated.bond_id']
```

	Scalar_Value
0	97.0387

📝 Summary:

1. The DataFrame contains a single column labeled `Scalar_Value` with a numerical value of 97.038739. This value represents a scalar measurement, likely the result of a calculation or aggregation from a larger dataset.
2. The `Query_Code` filters the original DataFrame to include only rows where the `activity_id` is 'run_individual_bde' and the `generated.bond_id` contains 'C-H'. It then calculates the mean of the `generated.bd_enthalpy` values for these filtered rows. The result in the reduced DataFrame, 97.038739, is the average bond dissociation enthalpy for bonds labeled 'C-H'.

The average bond dissociation enthalpy for bonds labeled 'C-H' in the current workflow is 97.04 kJ/mol.

Send a message





In this molecule, what is the multiplicity and charge of the parent molecule?



Here's the code:

```
result = df[df['activity_id'] == 'run_bde'][['generated.species.parent.multiplicity', 'generated.species.parent.charge']]
```

	generated.species.parent.multiplicity	generated.species.parent.charge
0	1	0
1	None	None

Summary:

1. The DataFrame contains two columns: `generated.species.parent.multiplicity` and `generated.species.parent.charge`. These columns represent the multiplicity and charge of the parent molecule, respectively. The first row shows values of 1.0 for multiplicity and 0.0 for charge, indicating a singlet state with a neutral charge. The second row contains NaN values, suggesting missing data for that particular entry.
2. The `Query_Code` specifically filters the DataFrame for entries with `activity_id` equal to 'run_bde' and retrieves the multiplicity and charge of the parent molecule. This indicates that the user is interested in understanding the electronic and ionic state of the parent molecule in a specific activity or experiment denoted by 'run_bde'.
3. The molecule has a multiplicity of 1 (singlet state) and a charge of 0 (neutral).

Send a message





In this molecule, what is the multiplicity and charge of any fragment molecule?



Here's the code:

```
result = df[['generated.multiplicity', 'generated.charge']].dropna()
```

	generated.multiplicity	generated.charge
0	1	0
1	1	0
2	2	0
3	2	0
4	2	0
5	2	0
6	2	0
7	2	0
8	2	0
9	2	0



Summary:

1. The DataFrame contains two columns: `generated.multiplicity` and `generated.charge`. All sampled rows show a multiplicity of 2.0 and a charge of 0.0.
2. The `Query_Code` specifically extracts these two columns and removes any rows with missing values (`dropna()`). This suggests that the query is designed to retrieve the multiplicity and charge of fragment molecules from a larger dataset, focusing only on entries that have complete information for these two attributes.

The fragment molecules in the molecule have a multiplicity of 2 and a charge of 0.

Send a message





In this molecule, what is the bond dissociation enthalpy value for the label that contains O-H?



Here's the code:

```
activity_id'] == 'run_individual_bde') & (df['generated.bd_enthalpy'].notna())
```

	used.frag.label	generated.bd_enthalpy
0	O-H_7	93.6152

Summary:

Analysis and Insights

The provided DataFrame contains two columns: `generated.bd_enthalpy` and `used.frag.label`. The `generated.bd_enthalpy` column appears to contain numerical values, likely representing bond dissociation enthalpy values. The `used.frag.label` column contains string values, which seem to be labels or identifiers for molecular fragments.

Query Code Analysis

The query code filters the original DataFrame to select rows where:

- `activity_id` is equal to `'run_individual_bde'`.
- `generated.bd_enthalpy` is not null (`notna()`).
- `used.frag.label` contains the string `'O-H'`.

This suggests that the query is searching for bond dissociation enthalpy values associated with molecular fragments containing the label `'O-H'`.

Answer to User Query

The bond dissociation enthalpy value for the label containing O-H is 93.6152.

Send a message

