1_TYK2_Structural_Data

April 8, 2022

[1]:
```
import pandas as pd
import seaborn as sns
```

[2]:
```
pdb_df = pd.read_csv('../data/tyk2_structures_plotting.csv')
pdb_df.shape
```

[2]:
```
(38, 3)
```

[3]:
```
pdb_df.head()
```

[3]:

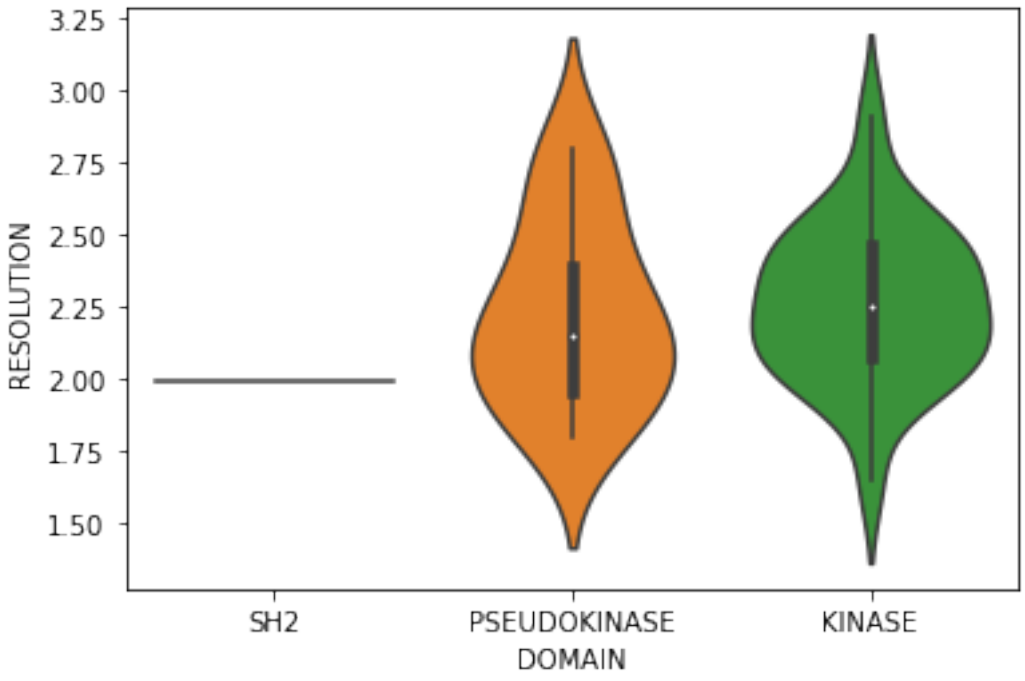
	PDB entry	Domain	Resolution ( )
0	4PO6	SH2	1.99
1	3ZON	PSEUDOKINASE	2.15
2	5C03	PSEUDOKINASE	1.90
3	5C01	PSEUDOKINASE	2.15
4	4OLI	PSEUDOKINASE	2.80

[4]:
```
pdb_df.columns = ['PDB_ID', 'DOMAIN', 'RESOLUTION']
pdb_df.head()
```

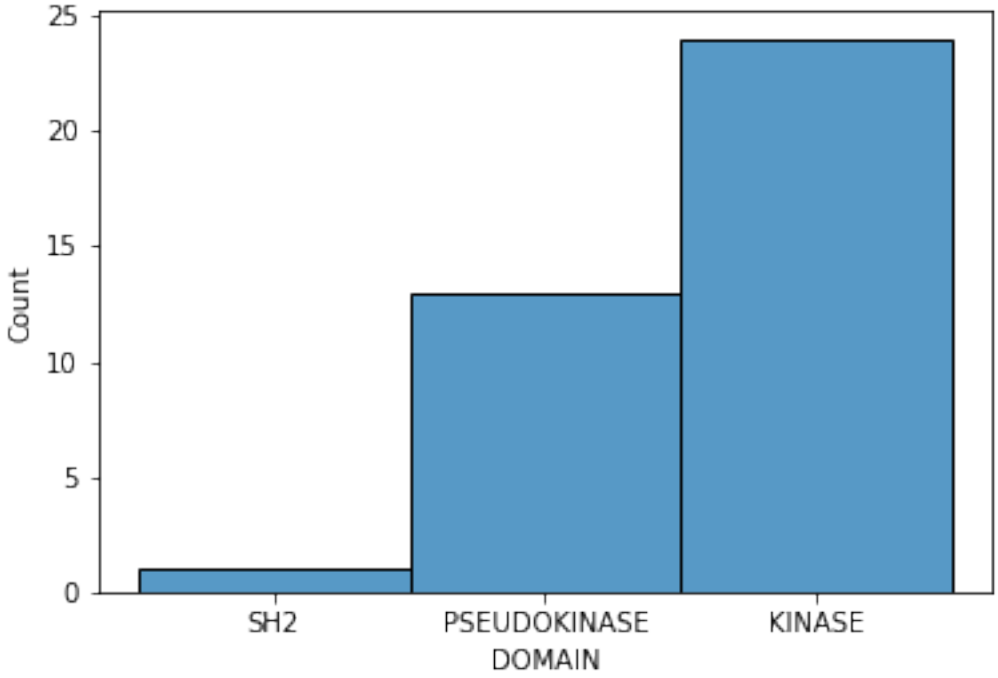
[4]:

	PDB_ID	DOMAIN	RESOLUTION
0	4PO6	SH2	1.99
1	3ZON	PSEUDOKINASE	2.15
2	5C03	PSEUDOKINASE	1.90
3	5C01	PSEUDOKINASE	2.15
4	4OLI	PSEUDOKINASE	2.80

[5]:
```
sns.violinplot(data=pdb_df, x=pdb_df['DOMAIN'], y=pdb_df['RESOLUTION']).figure.
↪savefig('../assets/tyk2_structures_violinplot.png')
```



[6]:
```
sns.histplot(data=pdb_df, x=pdb_df['DOMAIN']).figure.savefig('../assets/
↪tyk2_structures_histogram.png')
```



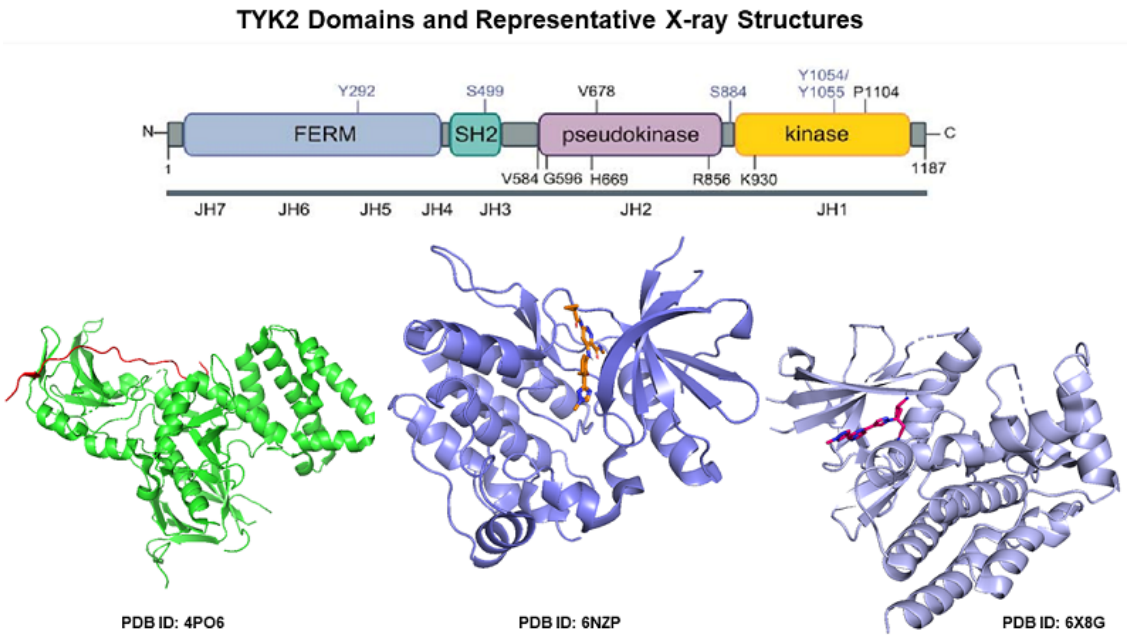
[7]:
```
pdb_df.describe()
```

[7]:

	RESOLUTION
count	38.000000
mean	2.243421
std	0.282472
min	1.650000
25%	2.007500
50%	2.170000
75%	2.437500
max	2.910000

[8]:
```
from collections import Counter
counts = list(Counter(pdb_df.DOMAIN).items())
counts
```

[8]:
```
[('SH2', 1), ('PSEUDOKINASE', 13), ('KINASE', 24)]
```



[ ]:
```
```