

# configuration\_analysis

August 26, 2018

**Analyse the distribution of side chain dihedral angles: Intro** The resultant cystine residue from the formation of the covalent disulfide bond between the thiol groups of two cysteine side chains is comprised of five dihedral angles. Based on X-ray structures we look at the distribution of these angles (X1, X2 and X3) to define their different conformations.

```
/home/david/coding/lib/python2.7/site-packages/ipykernel_launcher.py:9: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

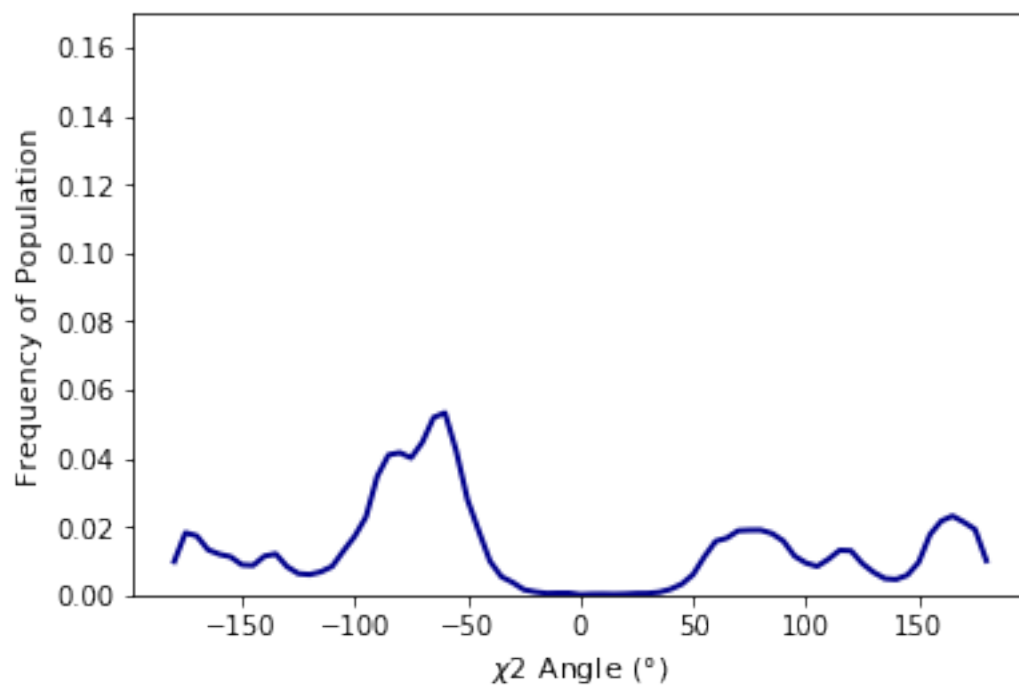
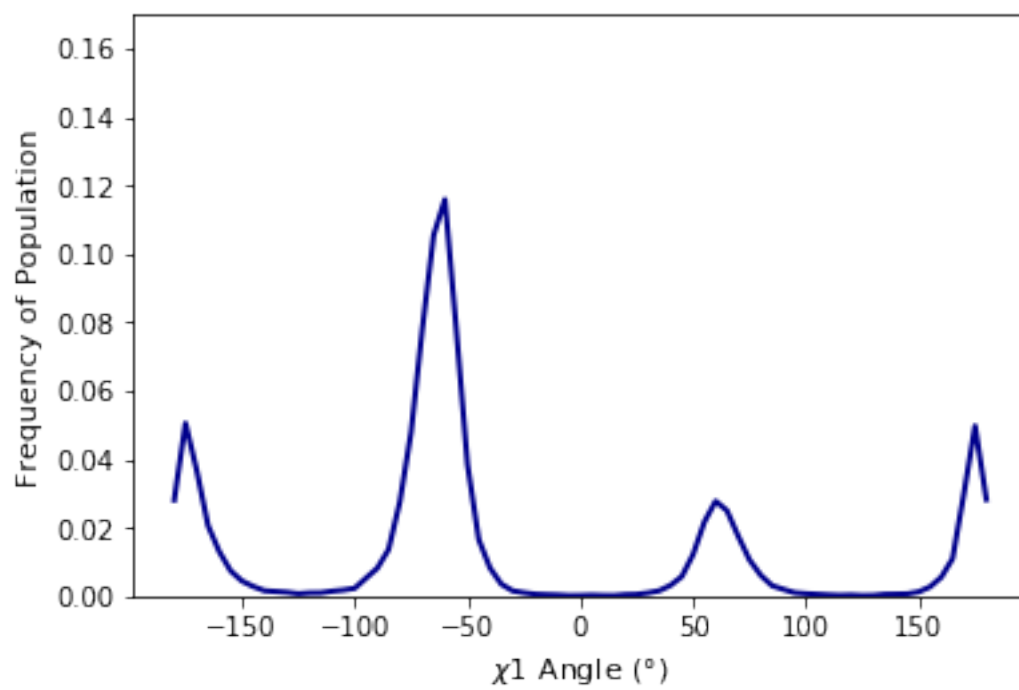
```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
    if __name__ == '__main__':
/home/david/coding/lib/python2.7/site-packages/ipykernel_launcher.py:10: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

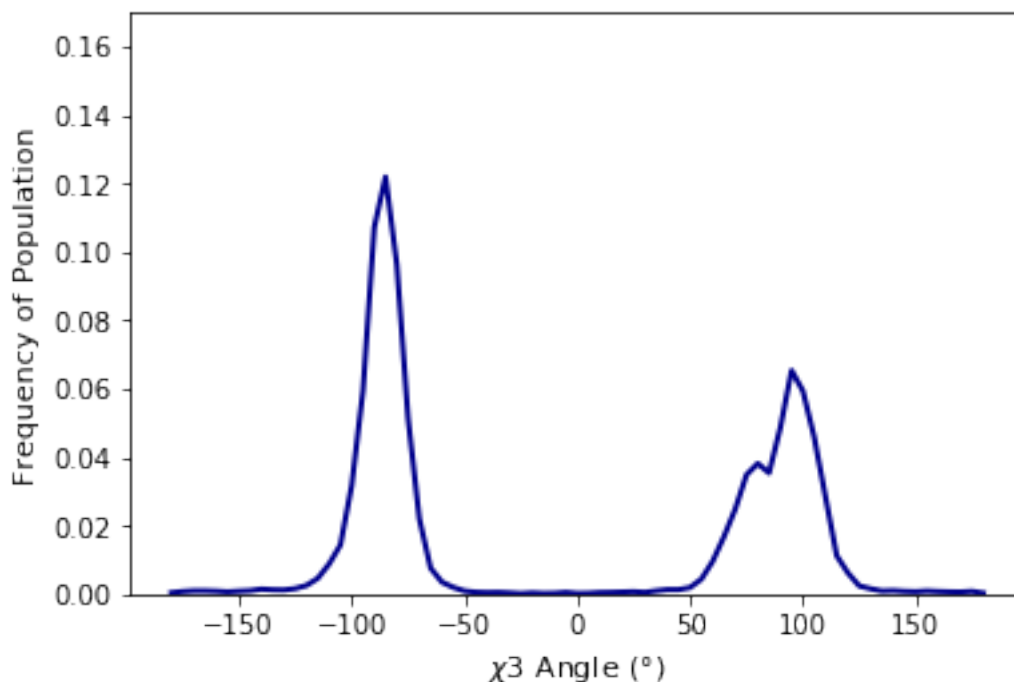
```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
    # Remove the CWD from sys.path while we load stuff.
/home/david/coding/lib/python2.7/site-packages/ipykernel_launcher.py:11: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
    # This is added back by InteractiveShellApp.init_path()
/home/david/coding/lib/python2.7/site-packages/ipykernel_launcher.py:12: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
    if sys.path[0] == '':
/home/david/coding/lib/python2.7/site-packages/ipykernel_launcher.py:13: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#
    del sys.path[0]
```





**Cystine dihedral distribution: Results** In agreement with previous analysis, can see that X1 adopts the 'Gauche-, Gauche+ and Trans' conformations with a very narrow distribution. Therefore defined as: Gauche+ (+60) = +30 to +90 Gauche- (-60) = -90 to -30 Trans (180) = 150 to 210

X2 also adopts these conformations, however a greater distribution is observed. Therefore the following ranges were used for classification Gauche+ (+60) = +30 to +120 Gauche- (-60) = -120 to -30 Trans (180) = 150 to 210

X3 angle distribution follows the previously identified 'Right Handed (+90)' or 'Left Handed (-90)' Right Handed (+90) = +60 to +120 Left Handed (-90) = -120 to -60

**Frequency of Configurations: Intro** Based on the definitions above, there is a possible 90 configurations a cystine can adopt when considering all five X angles (X1, X2, X3, X2, X1). Below calculates the frequency of those configurations If a disulfide had a dihedral angle out of the defined ranges above it was excluded from analysis

Total Number of Disulfides: 19605

Total Number of Disulfides in Defined Configurations for Structural Analysis: 15538

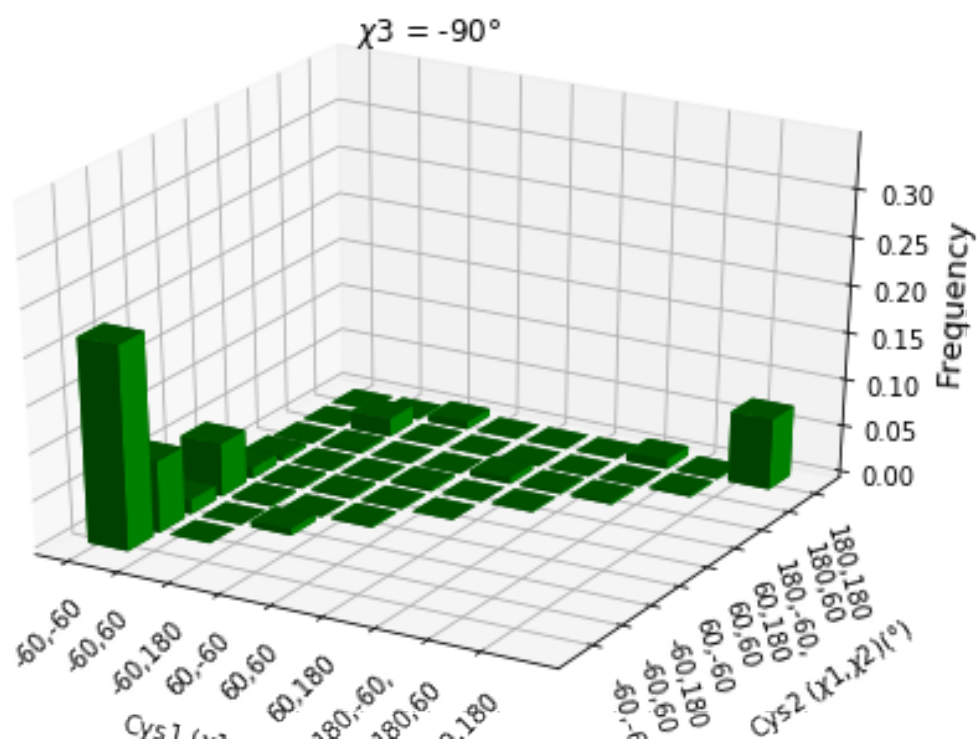
```
/home/david/coding/lib/python2.7/site-packages/ipykernel_launcher.py:192: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

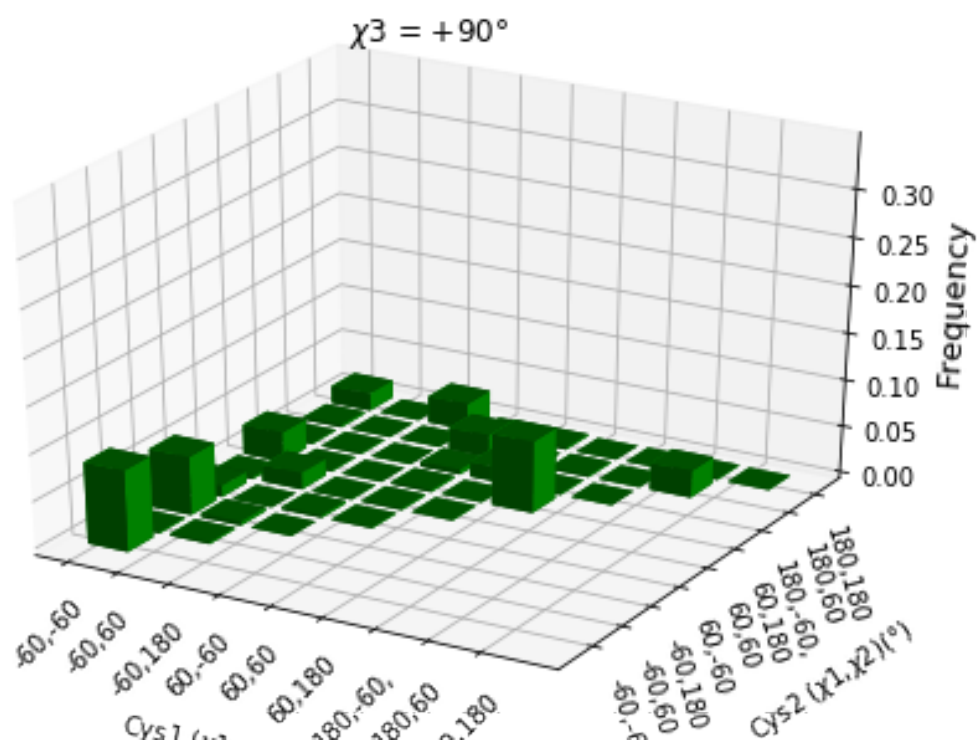
See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#>

**Generate Configuration Frequency Graphs** Generate a 3D plot tha shows the frequency of pop-  
ulation for each configuration

## 1 Generate configuration graphs for all cystines

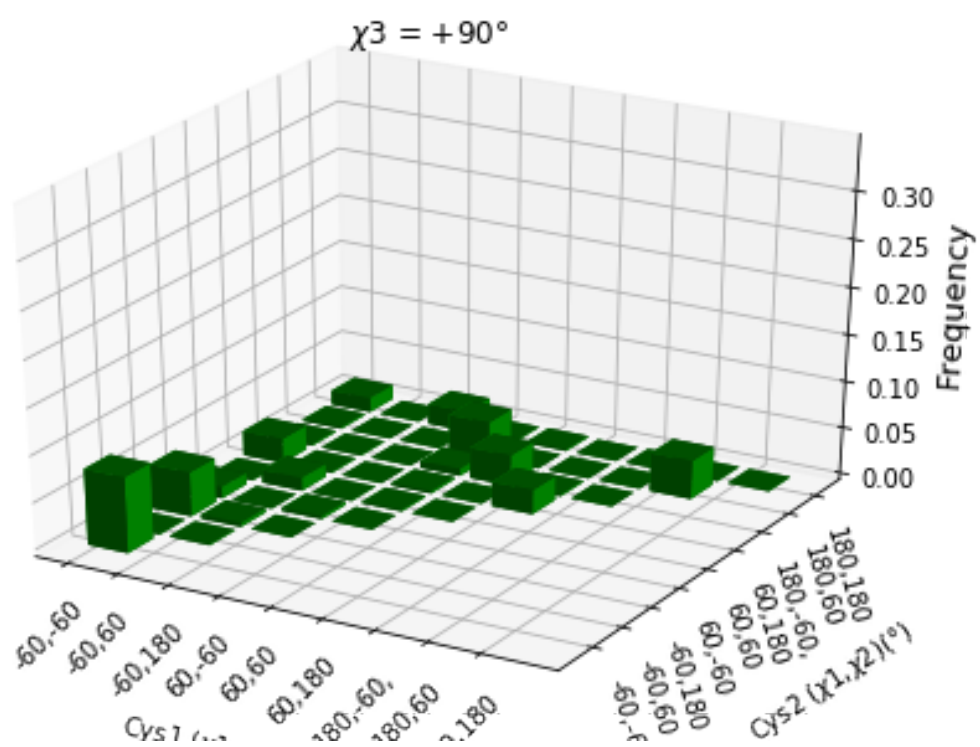
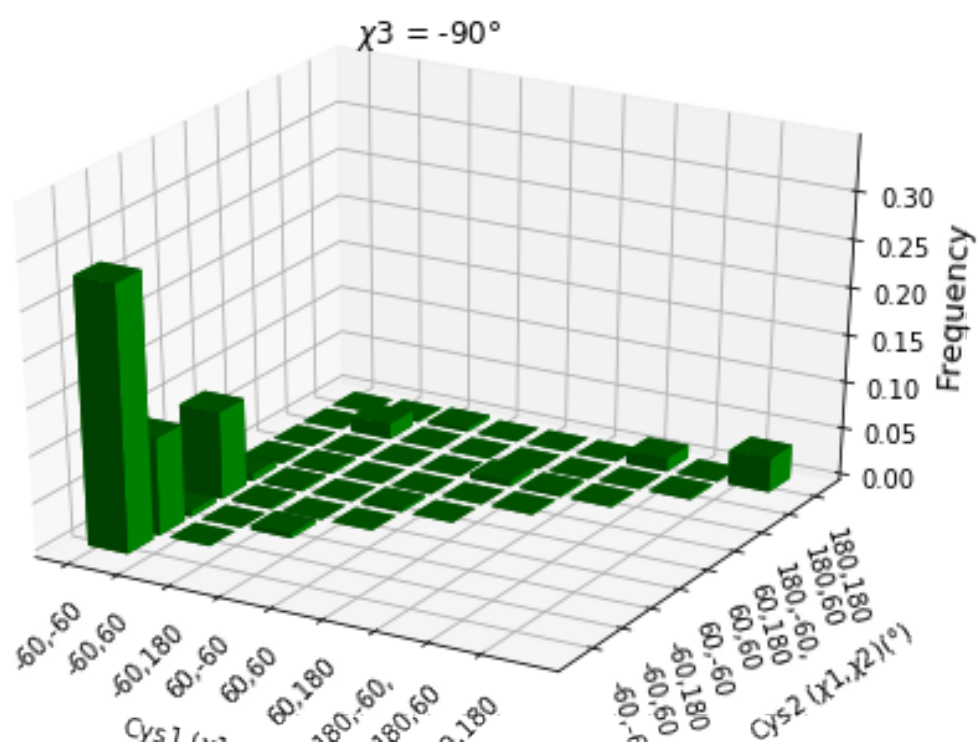
Out [5] : ( )





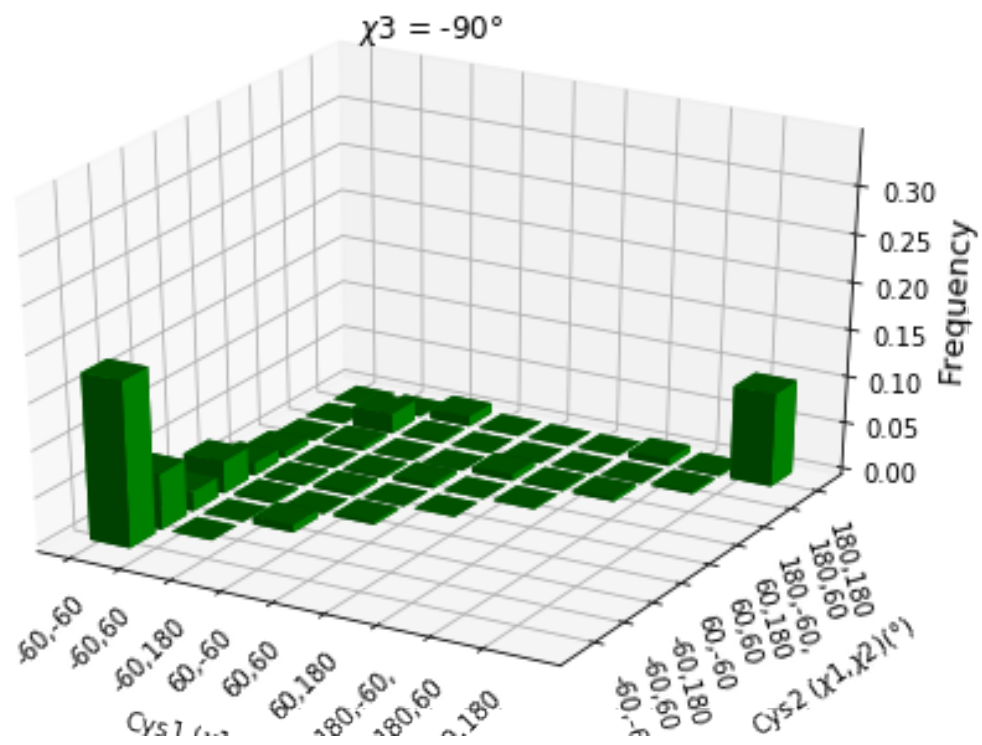
## 1.1 Generate configuration graphs for peptide cystines

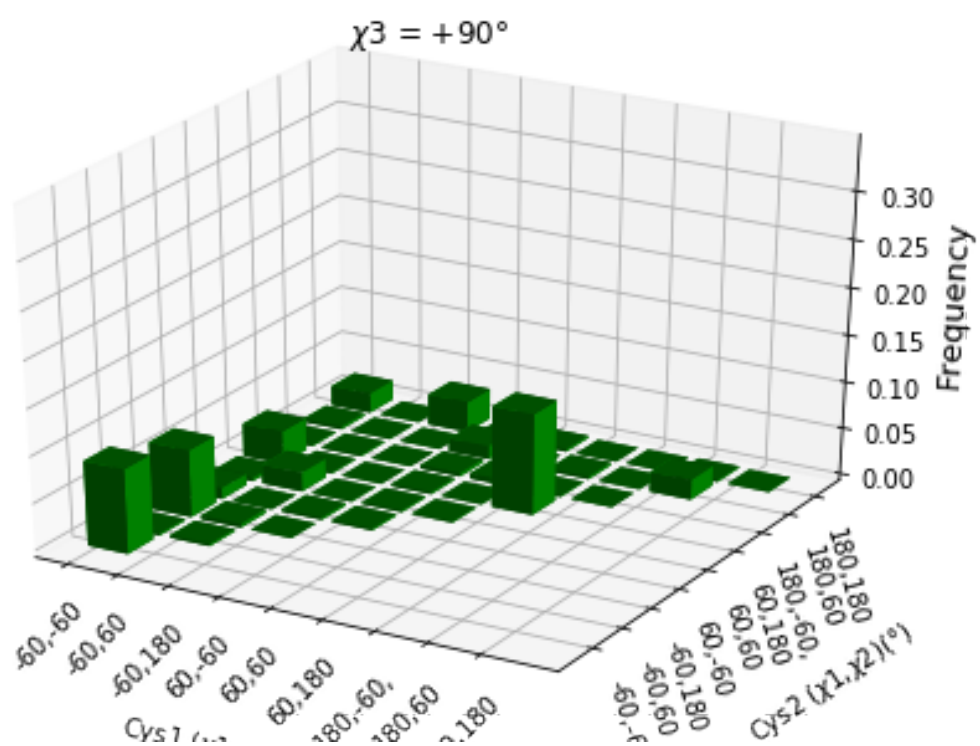
Out[6]: ()



## 1.2 Generate configuration graphs for protein cystines

Out[7]: ()

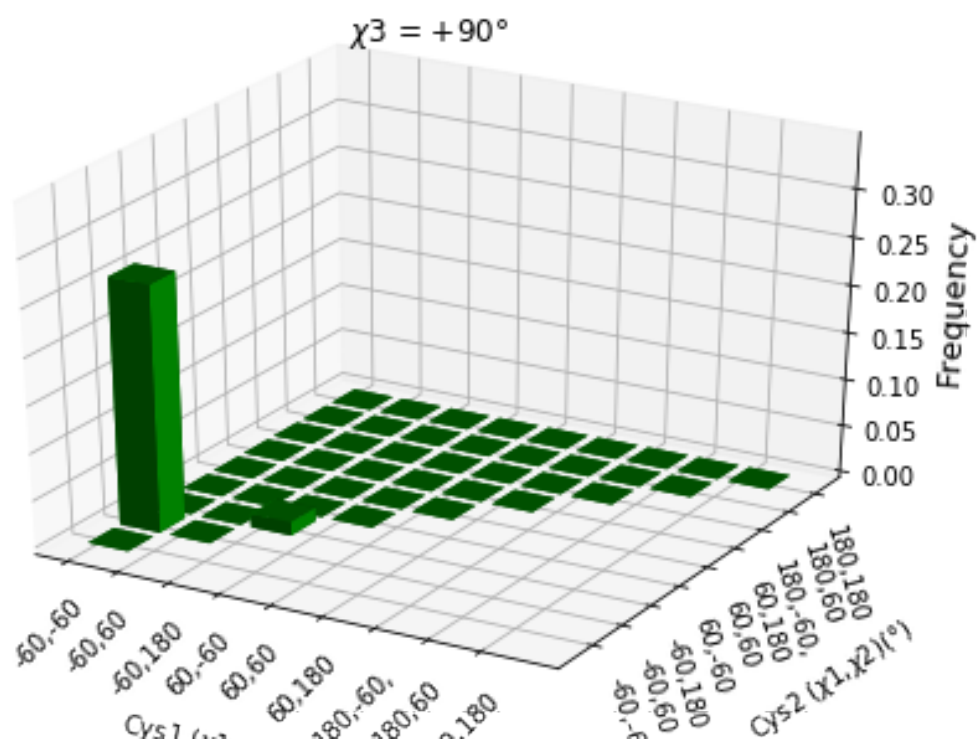
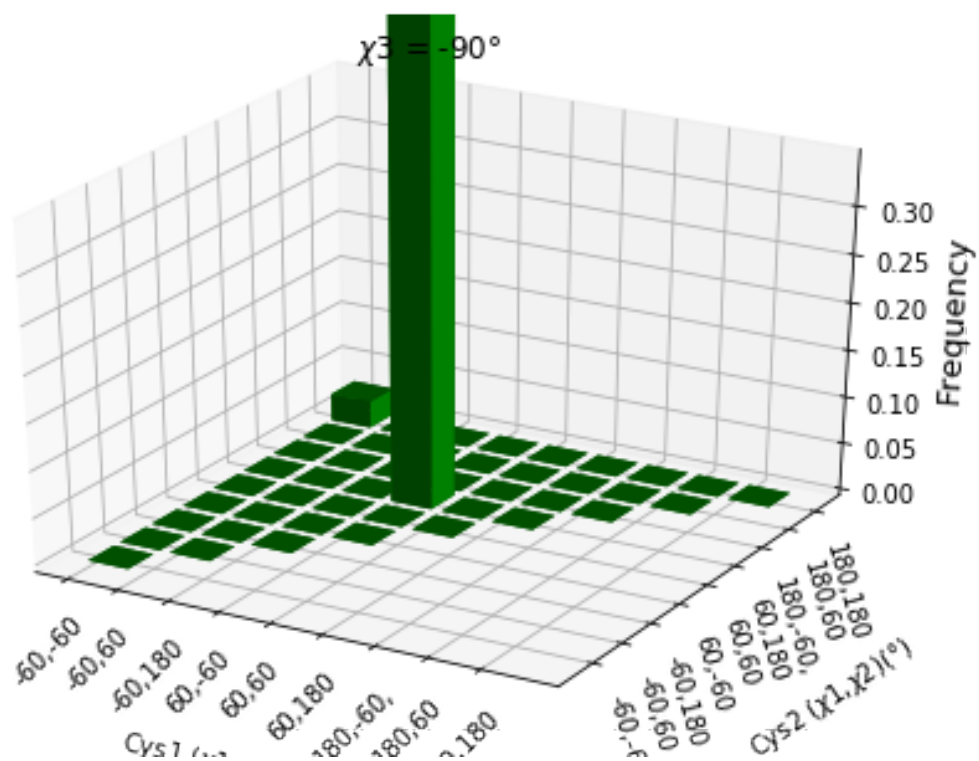




### 1.3 For Vicinal Disulfides

Out[8]: ()





## 1.4 CONFIGURATION FUNCTION

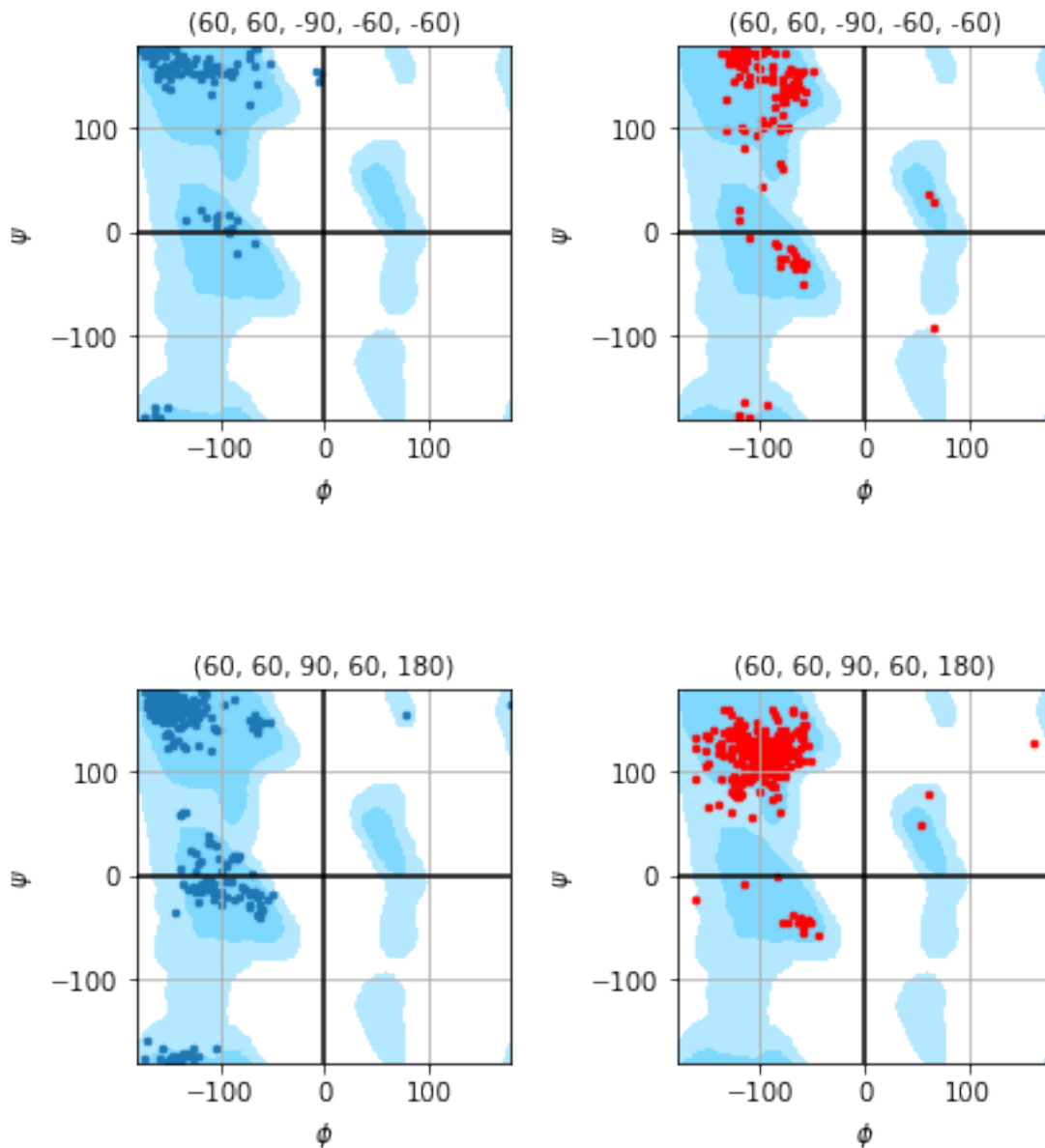
1.4.1 The below function returns individual dataframes for each configuration

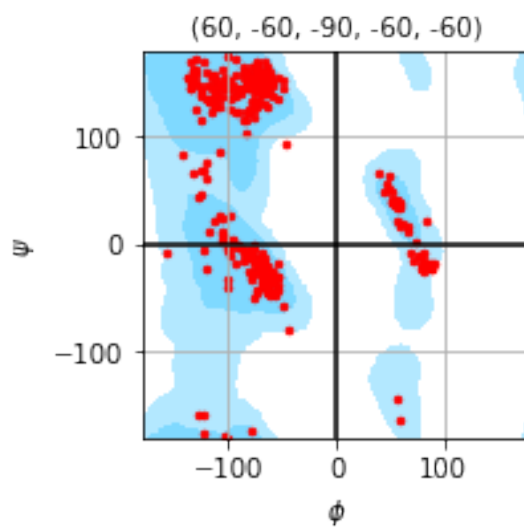
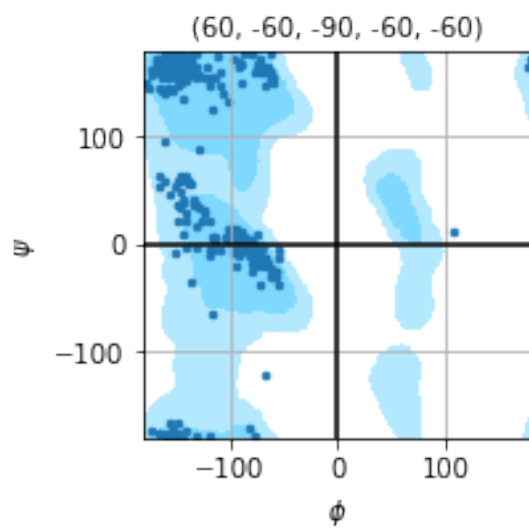
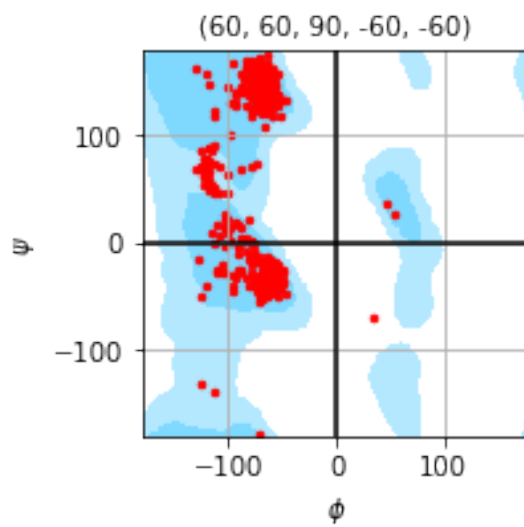
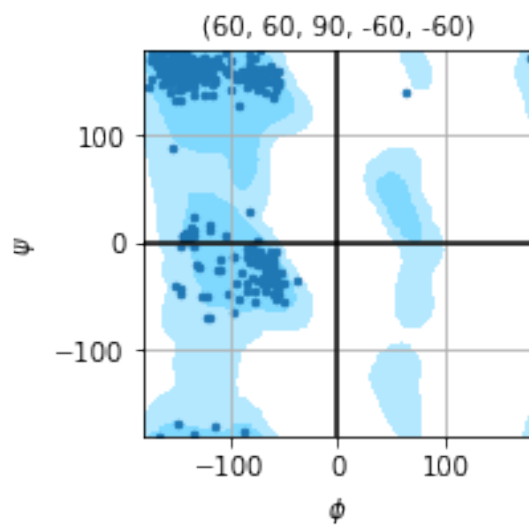
1.4.2 It will be called to generate figures that investigate structural features based on configuration

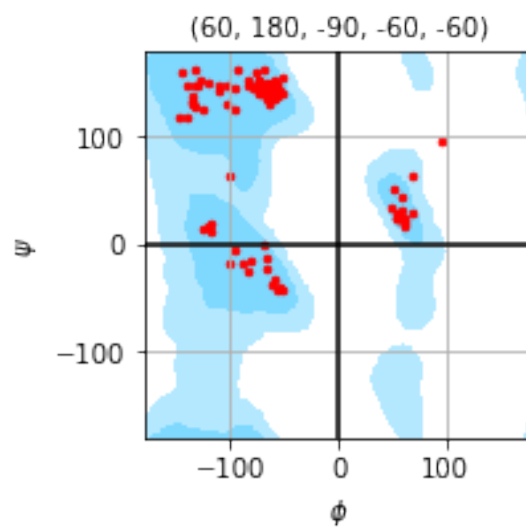
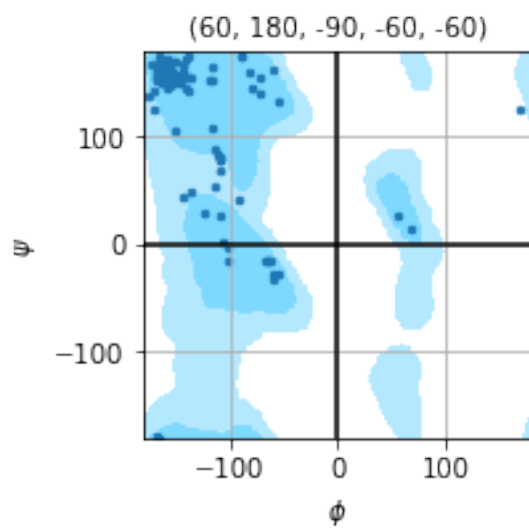
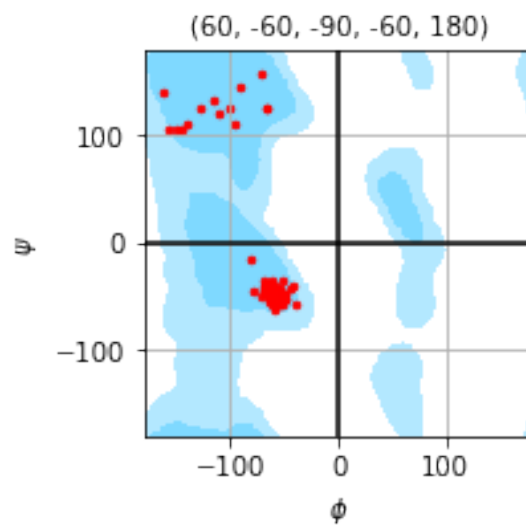
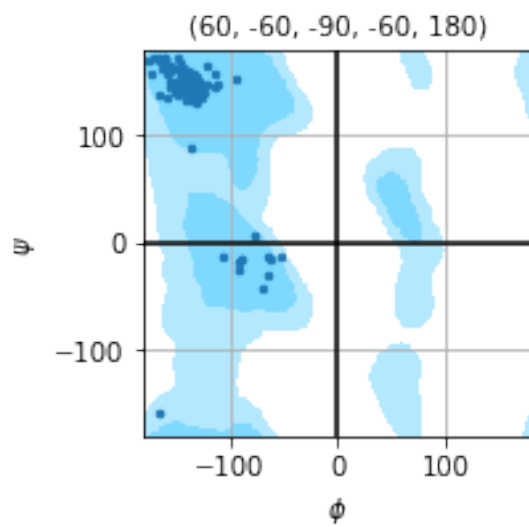
1.4.3 It re-writes the dataframe to ensure  $\text{Cys1} - \text{Cys2} == \text{Cys2} - \text{Cys1}$  based on configuration order

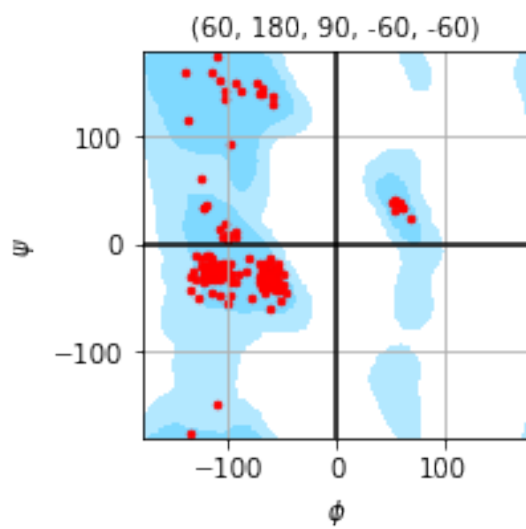
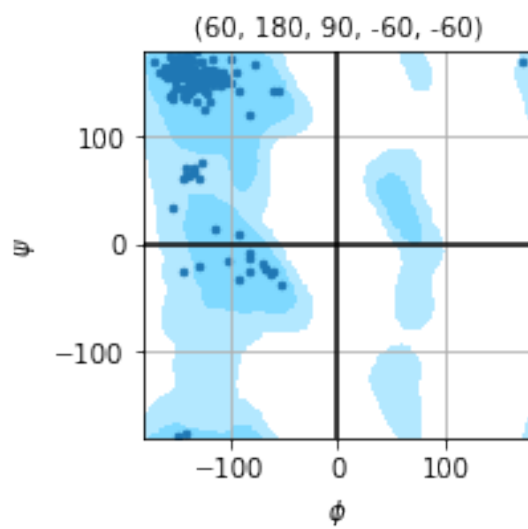
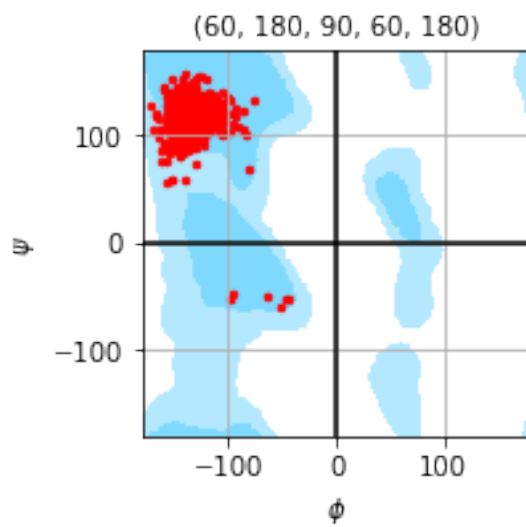
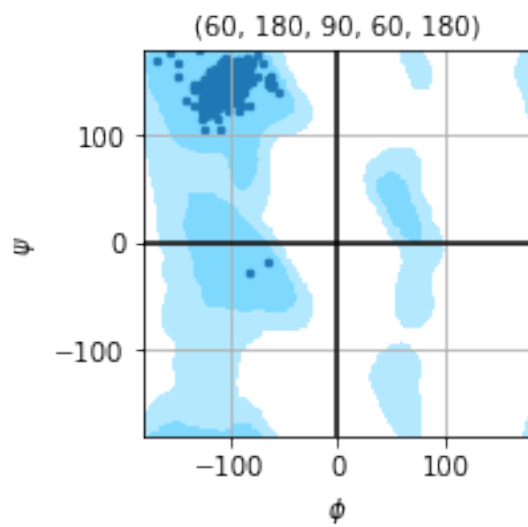
1.5 Function to generate Ramachandran Plots

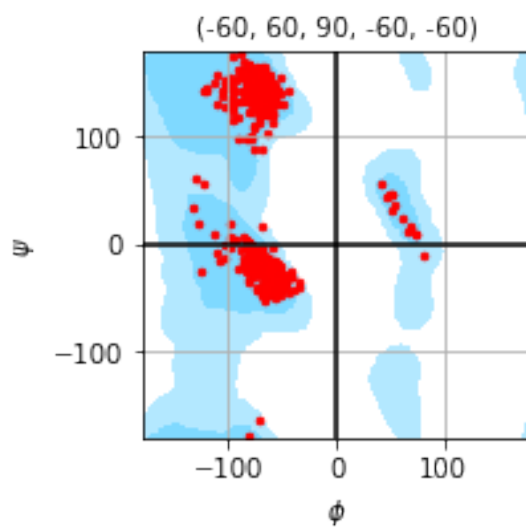
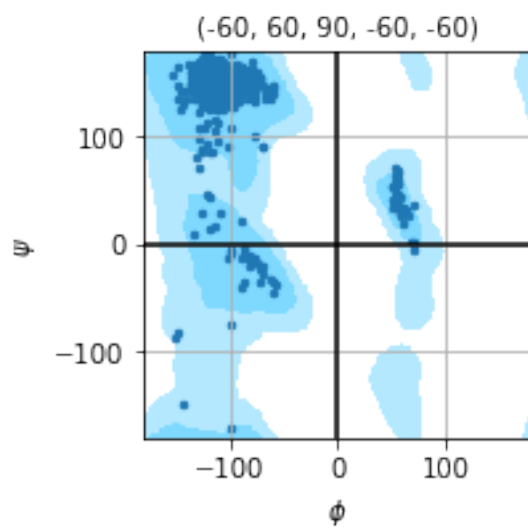
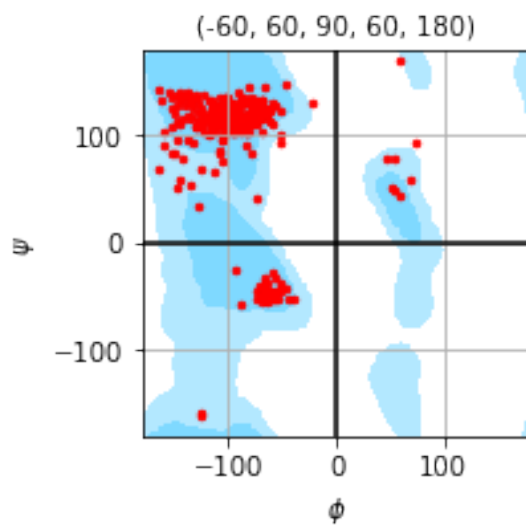
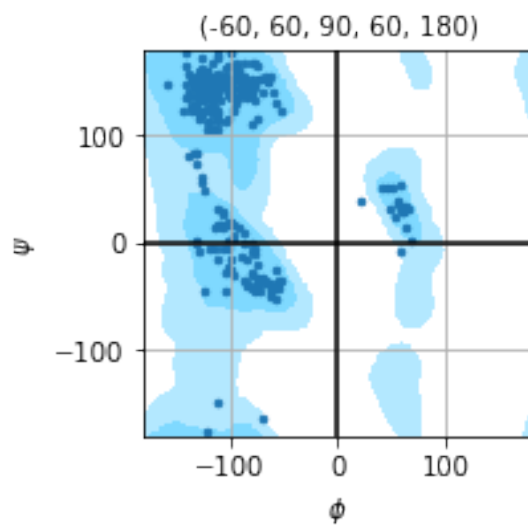
1.6 Call each configuration and

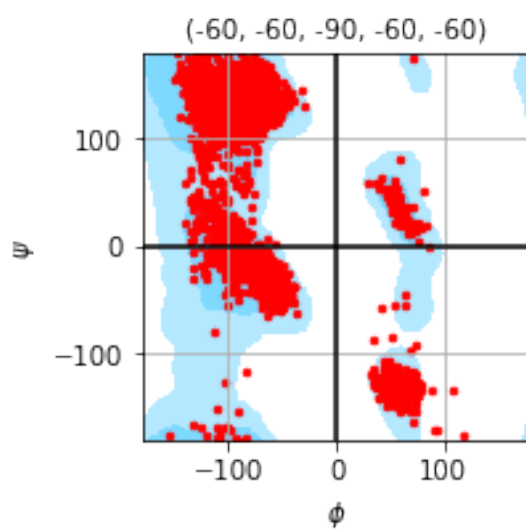
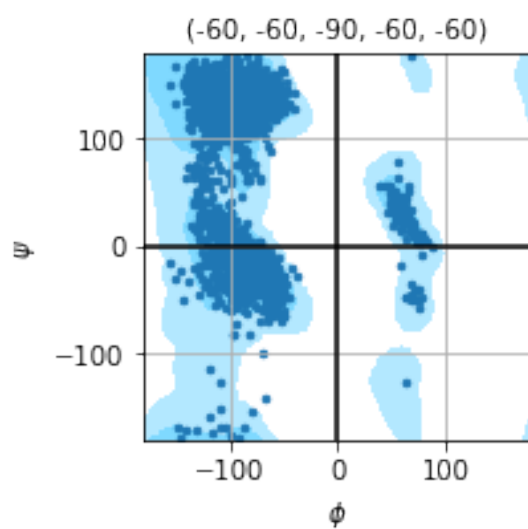
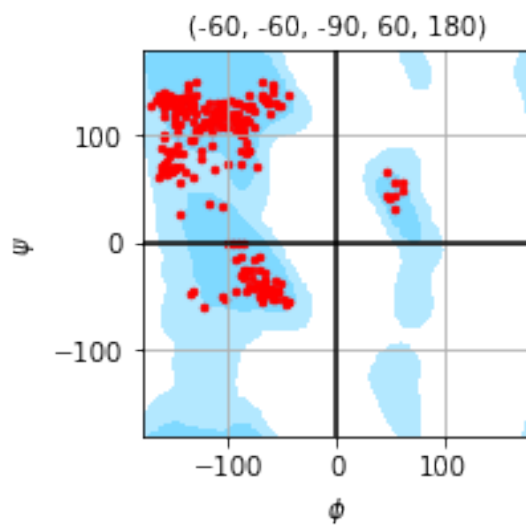
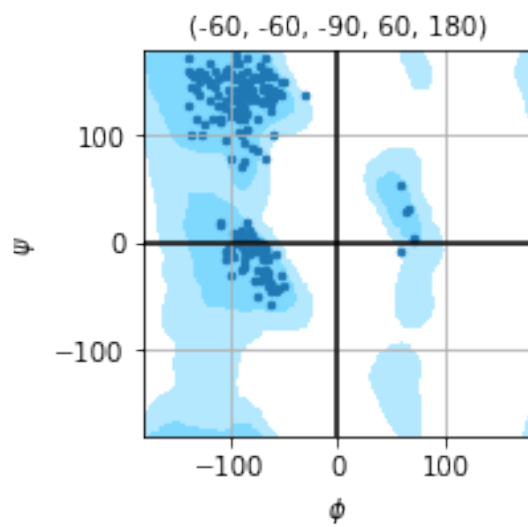


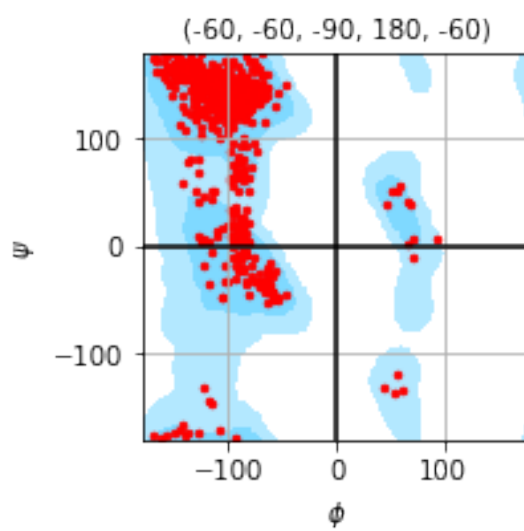
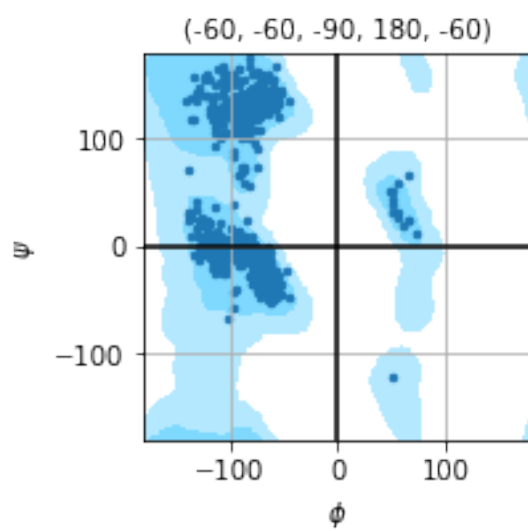
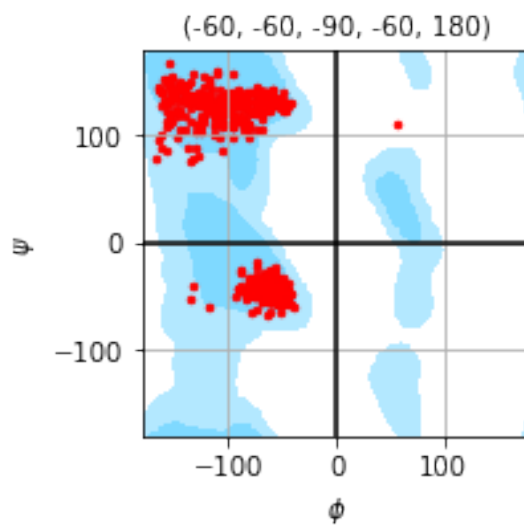
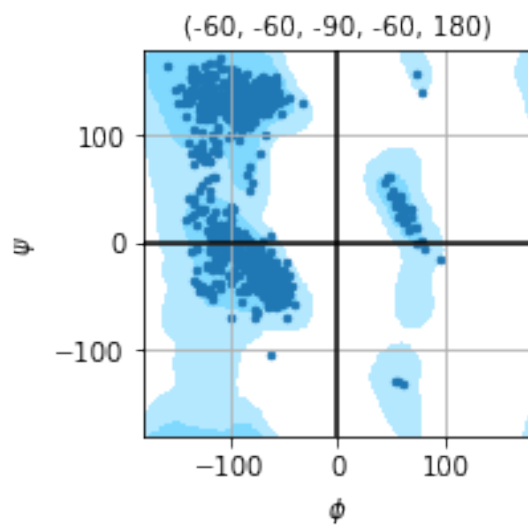




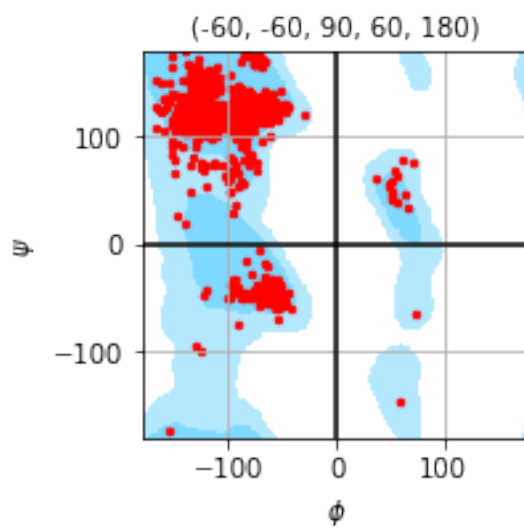
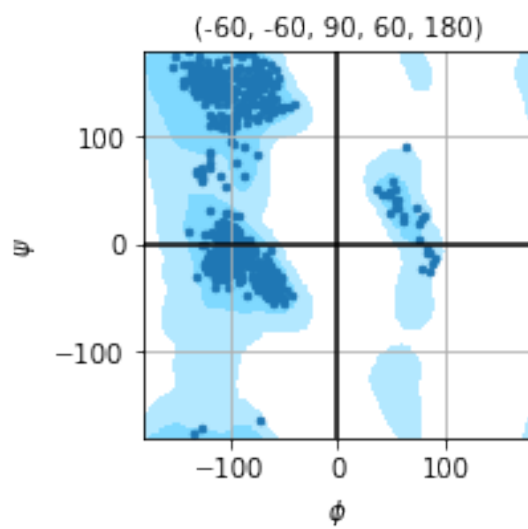
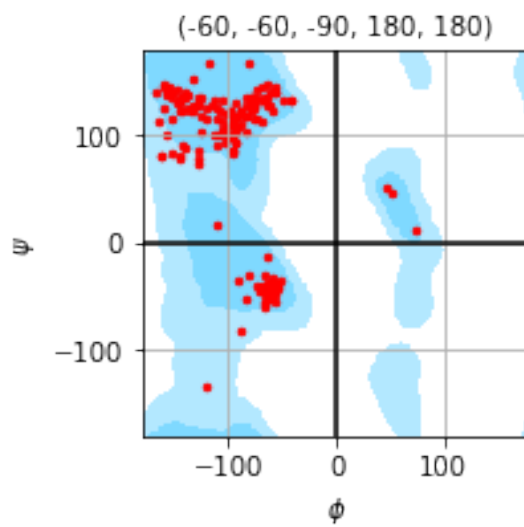
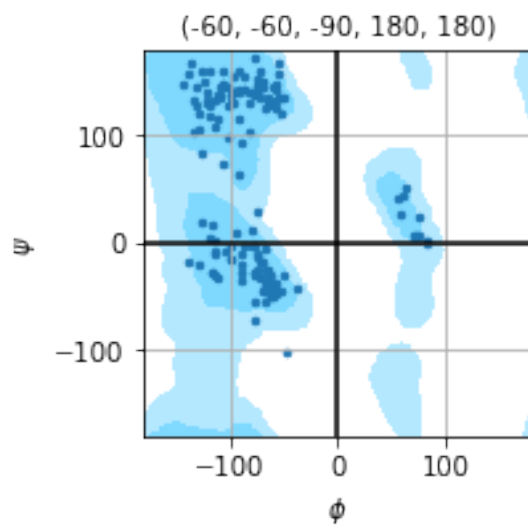


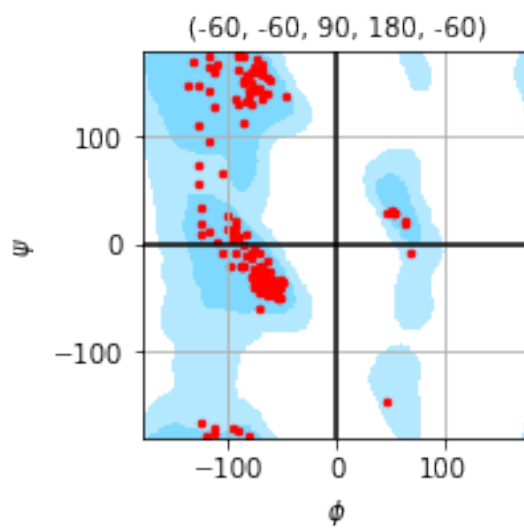
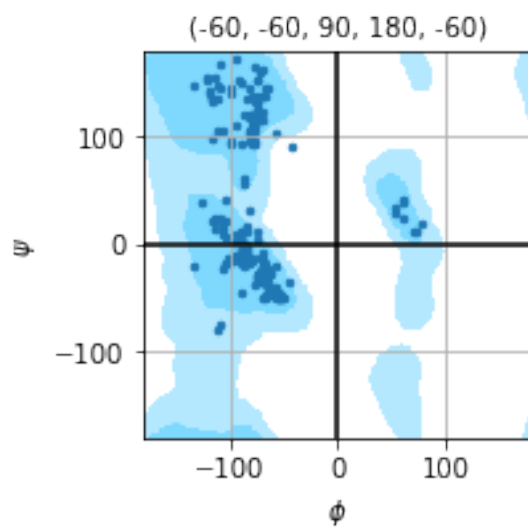
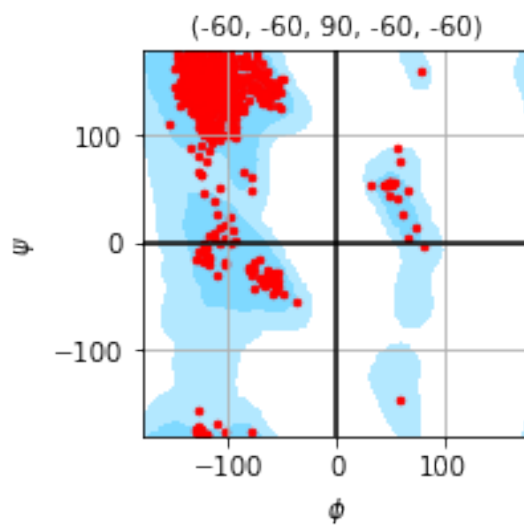
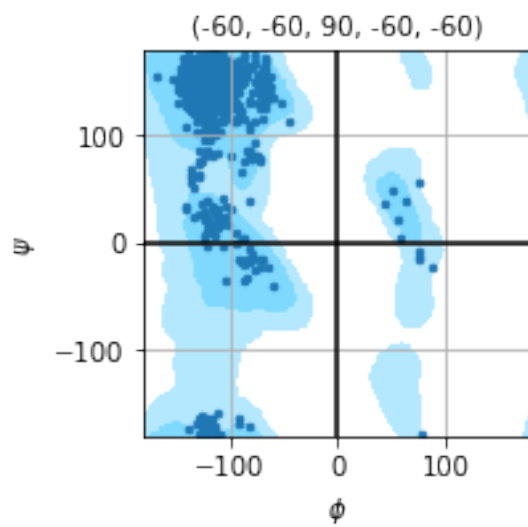


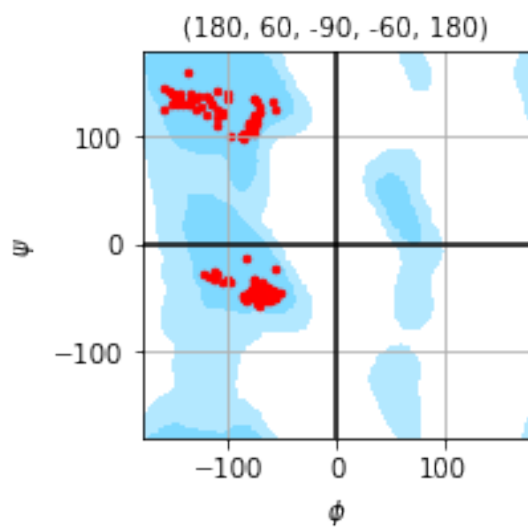
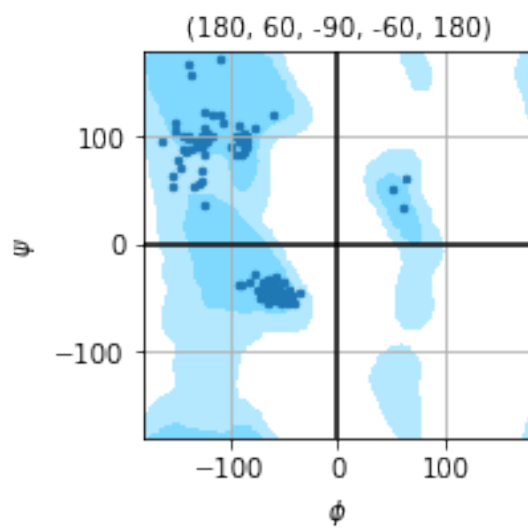
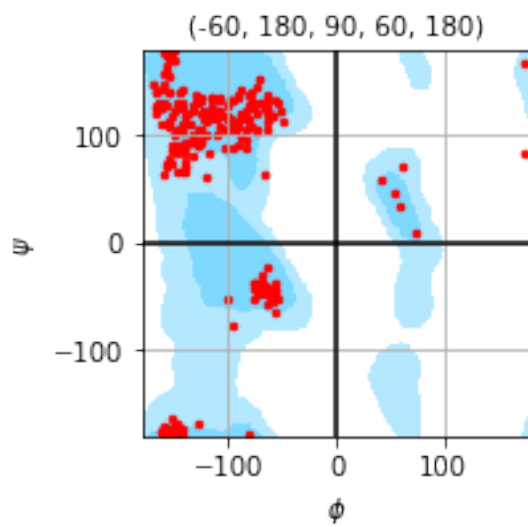
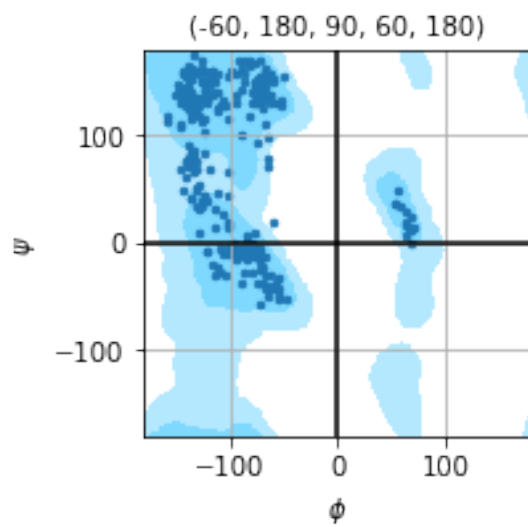


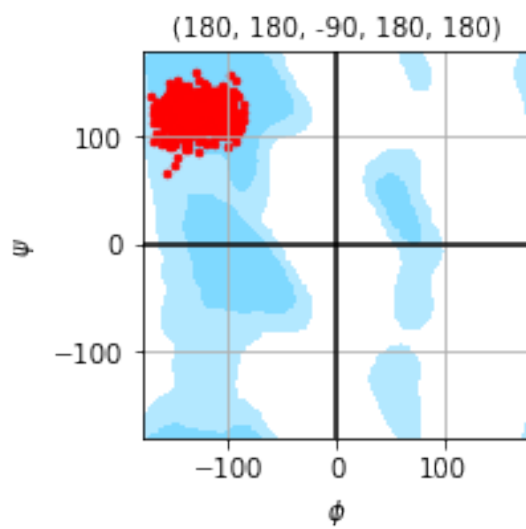
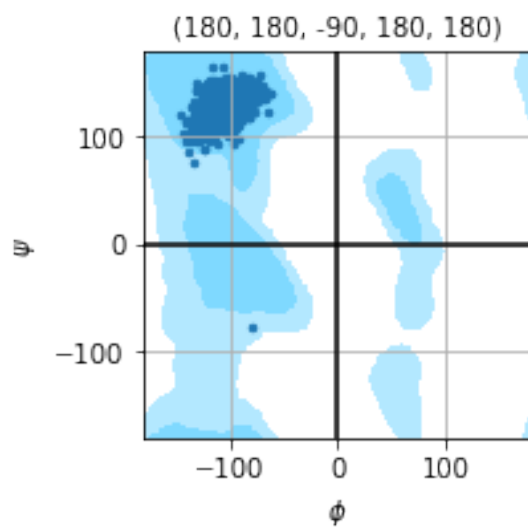
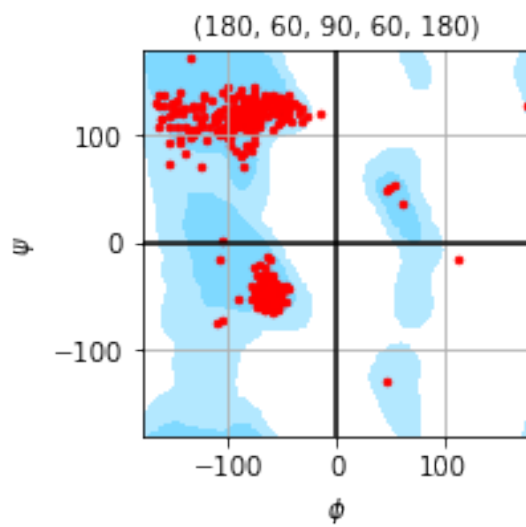
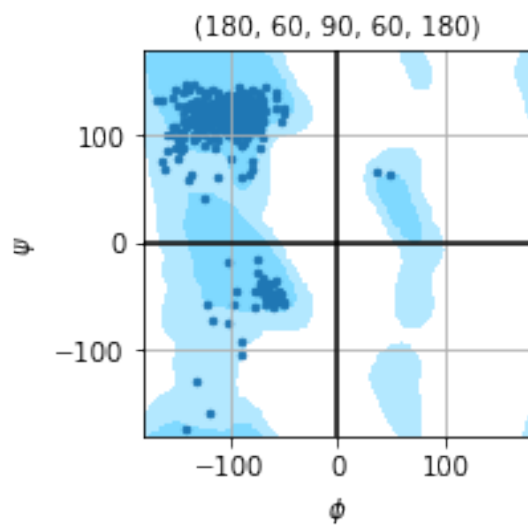


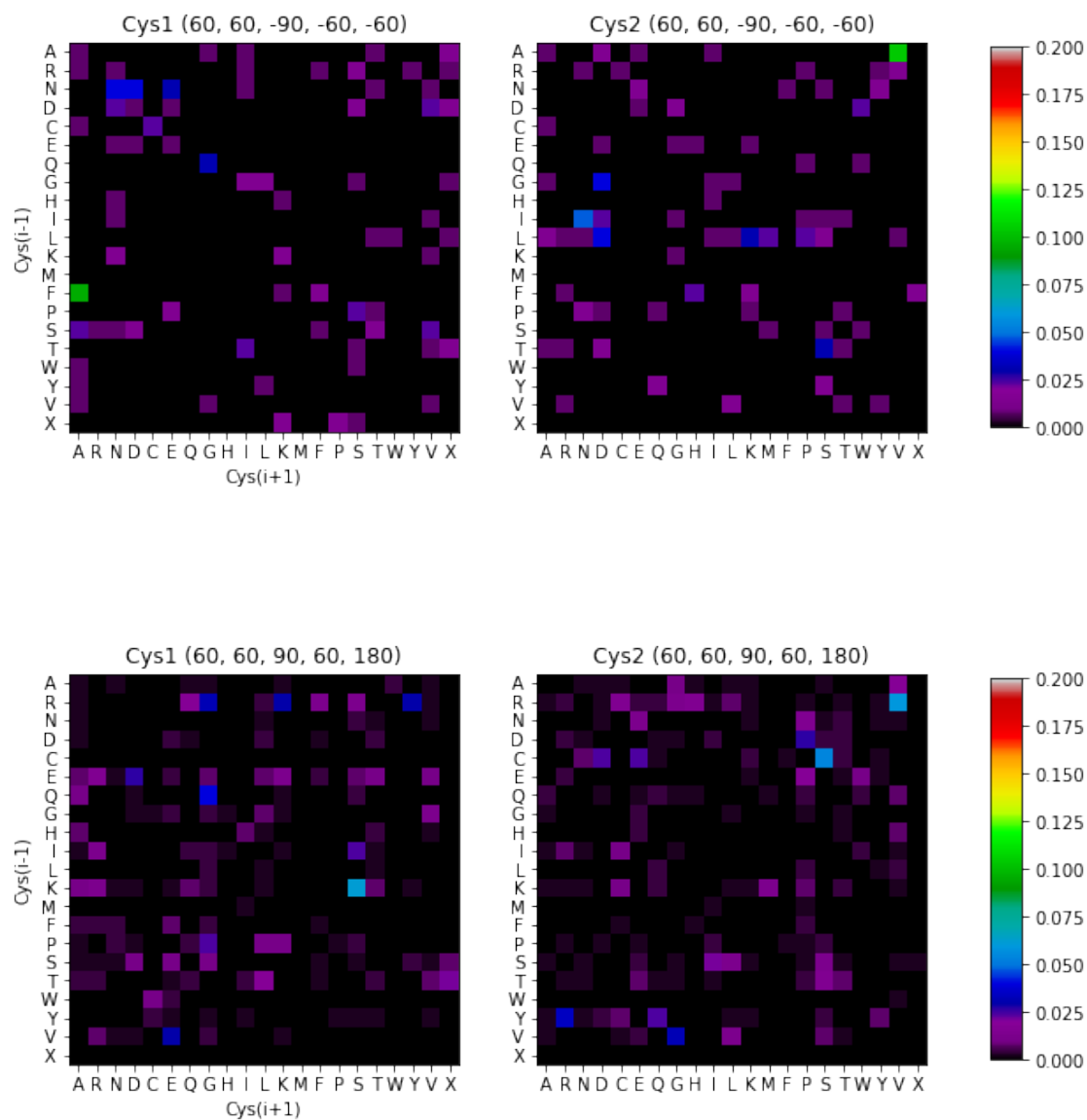


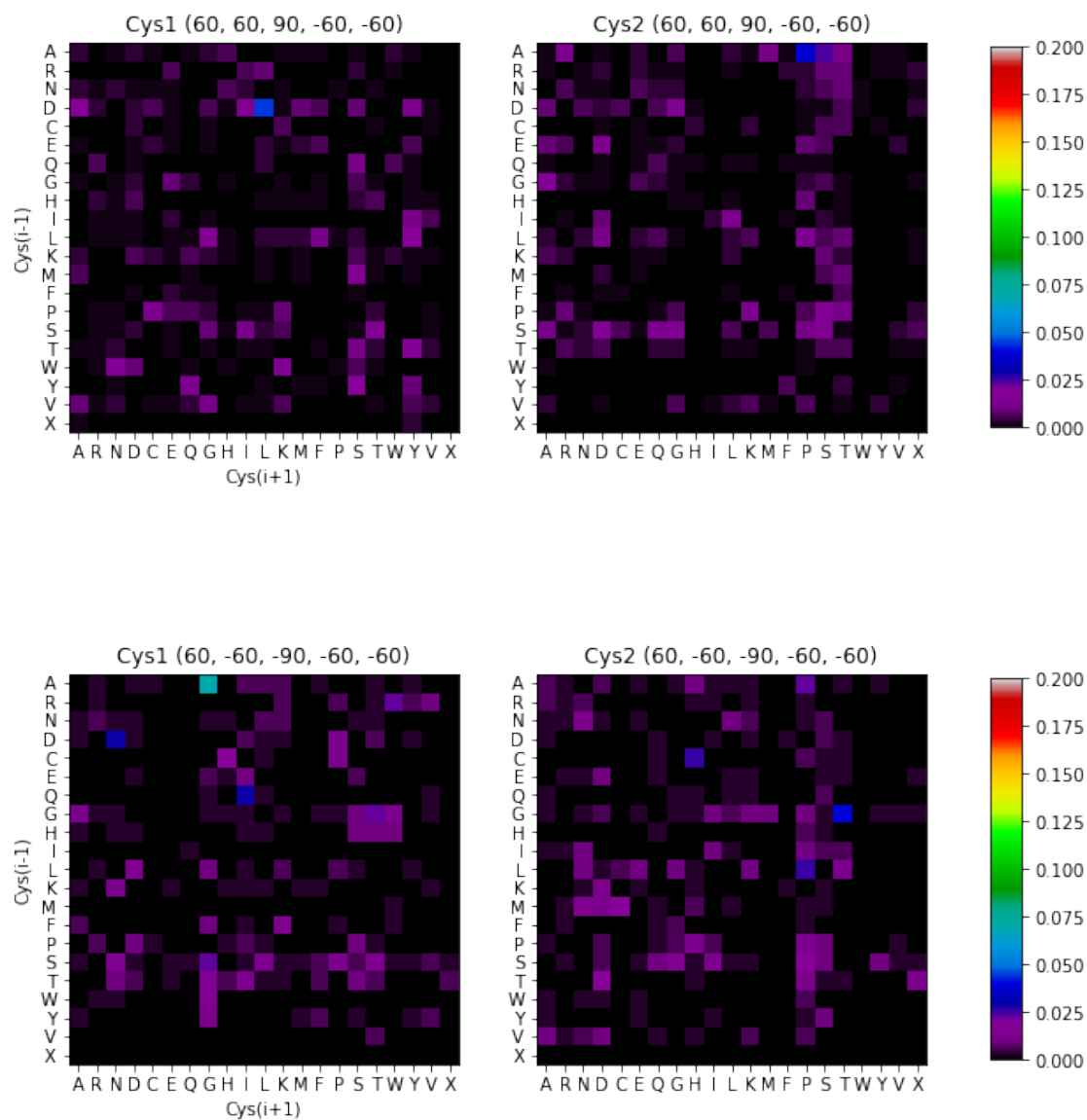


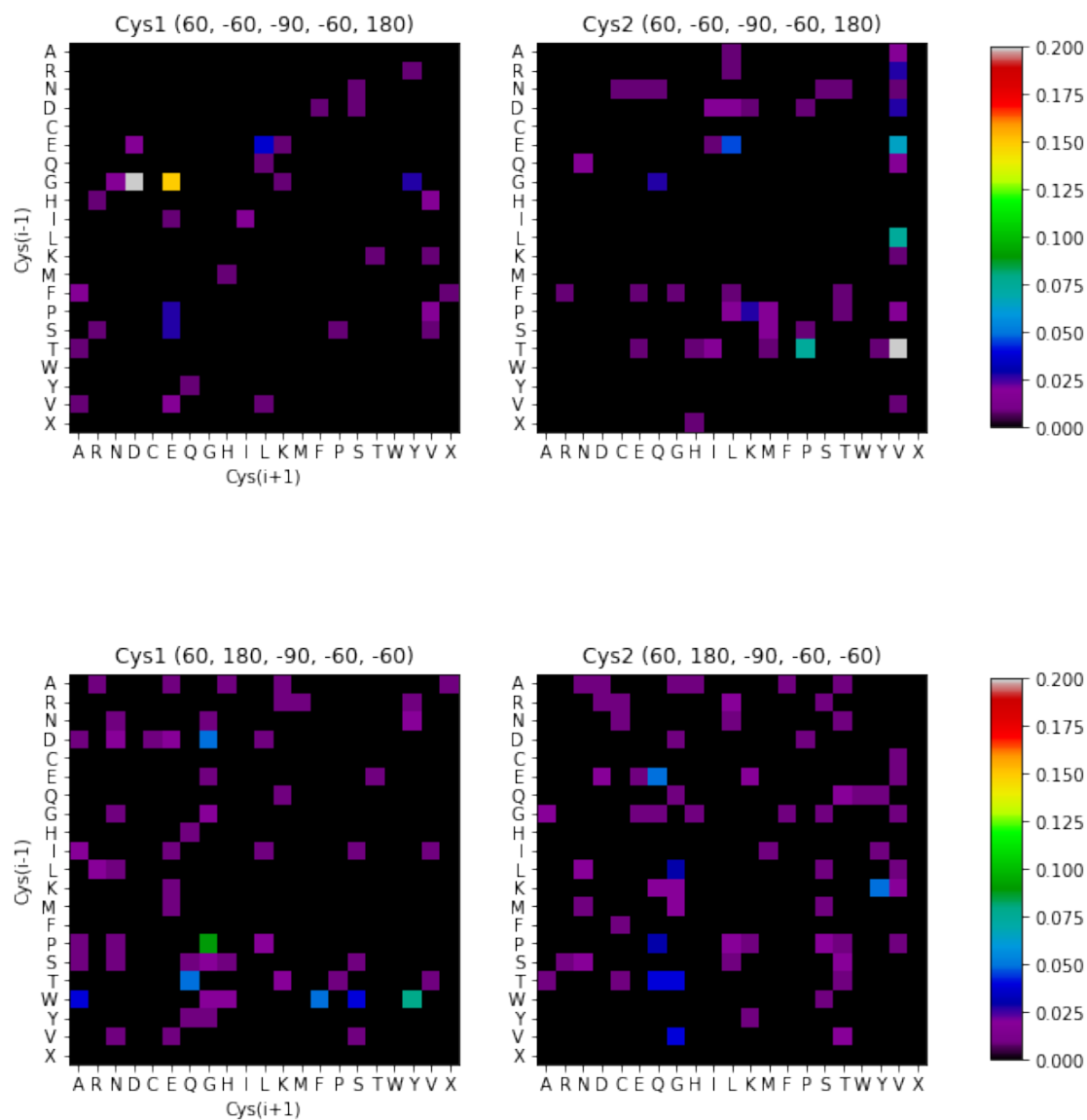


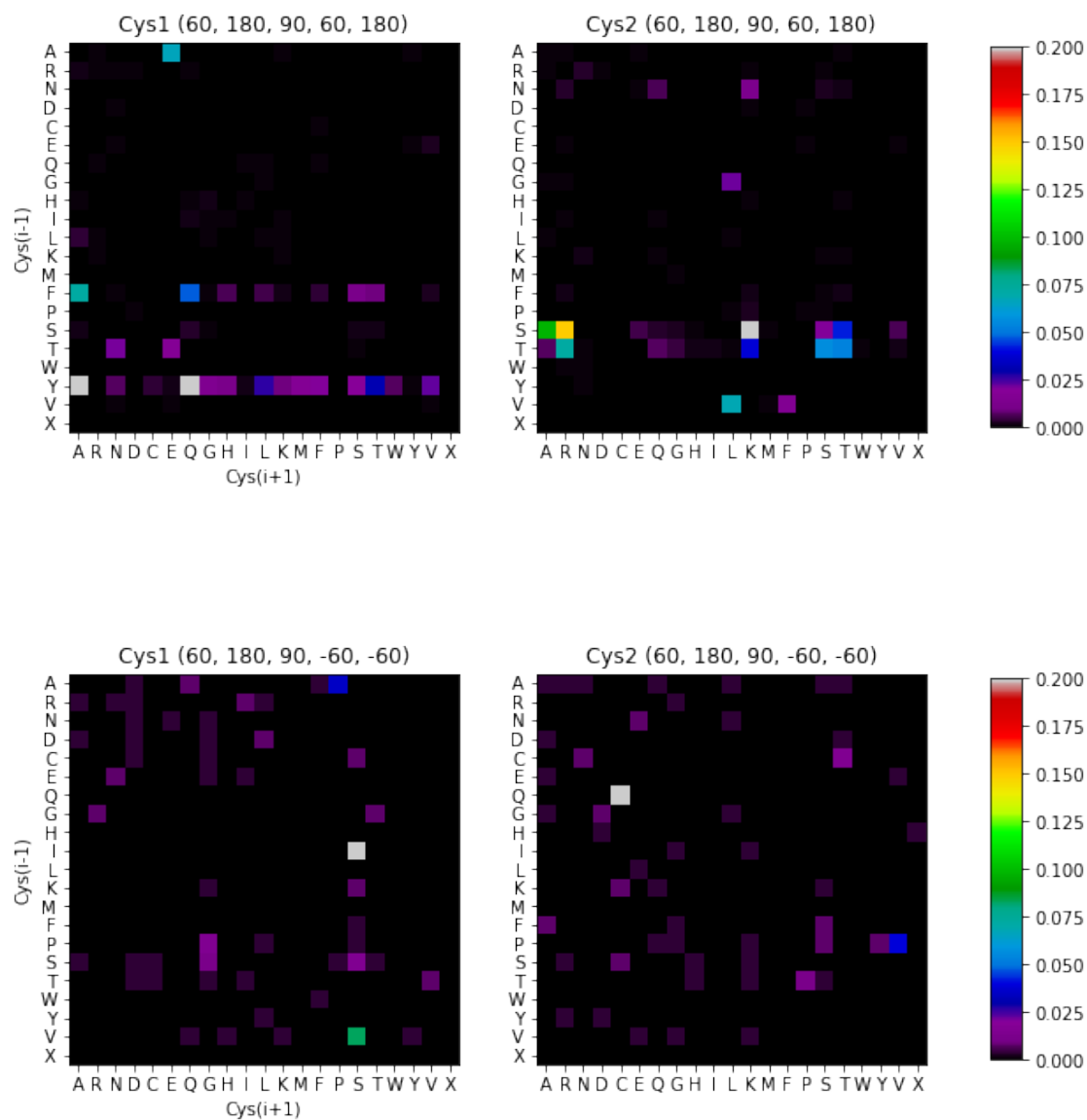




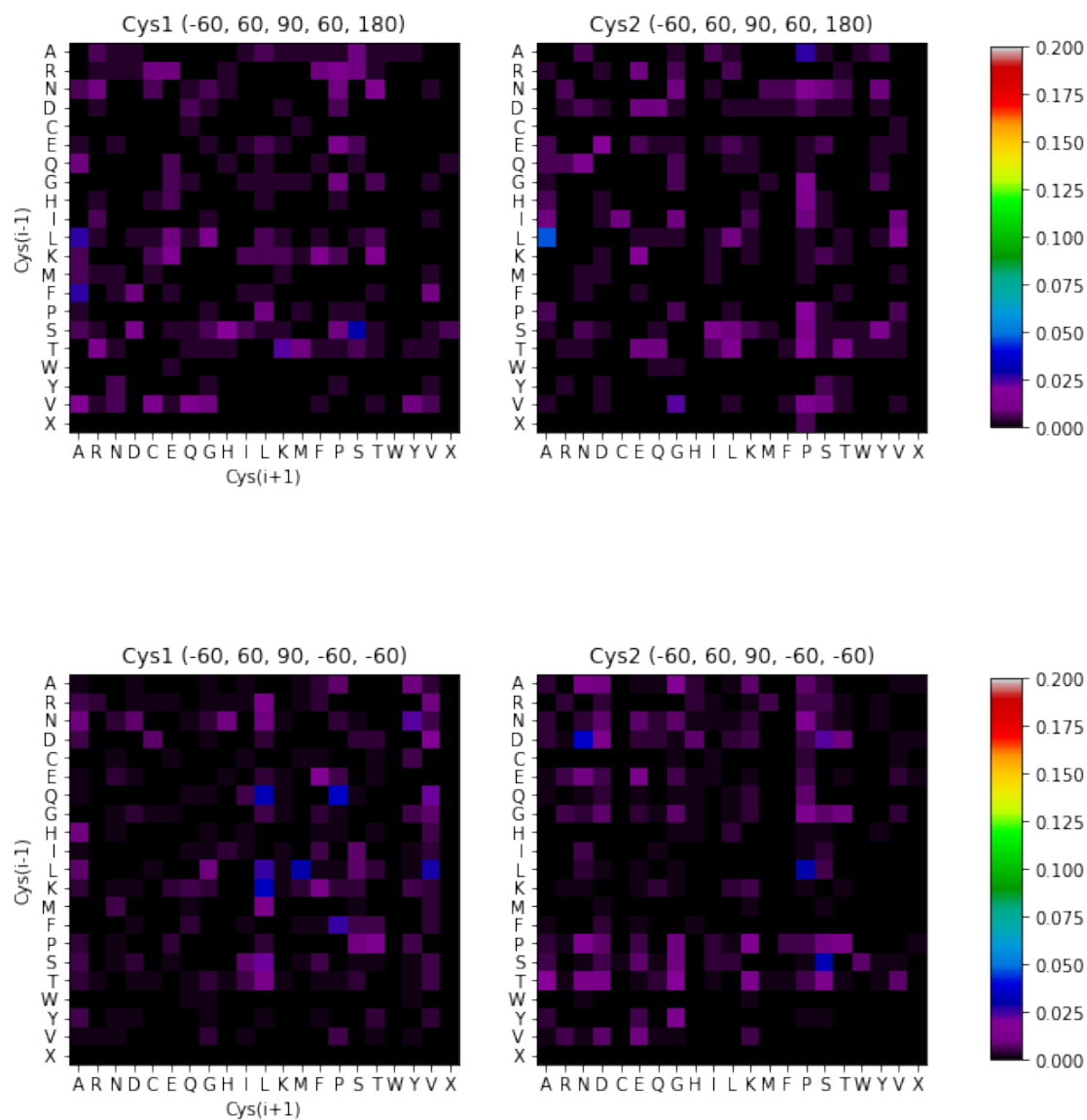


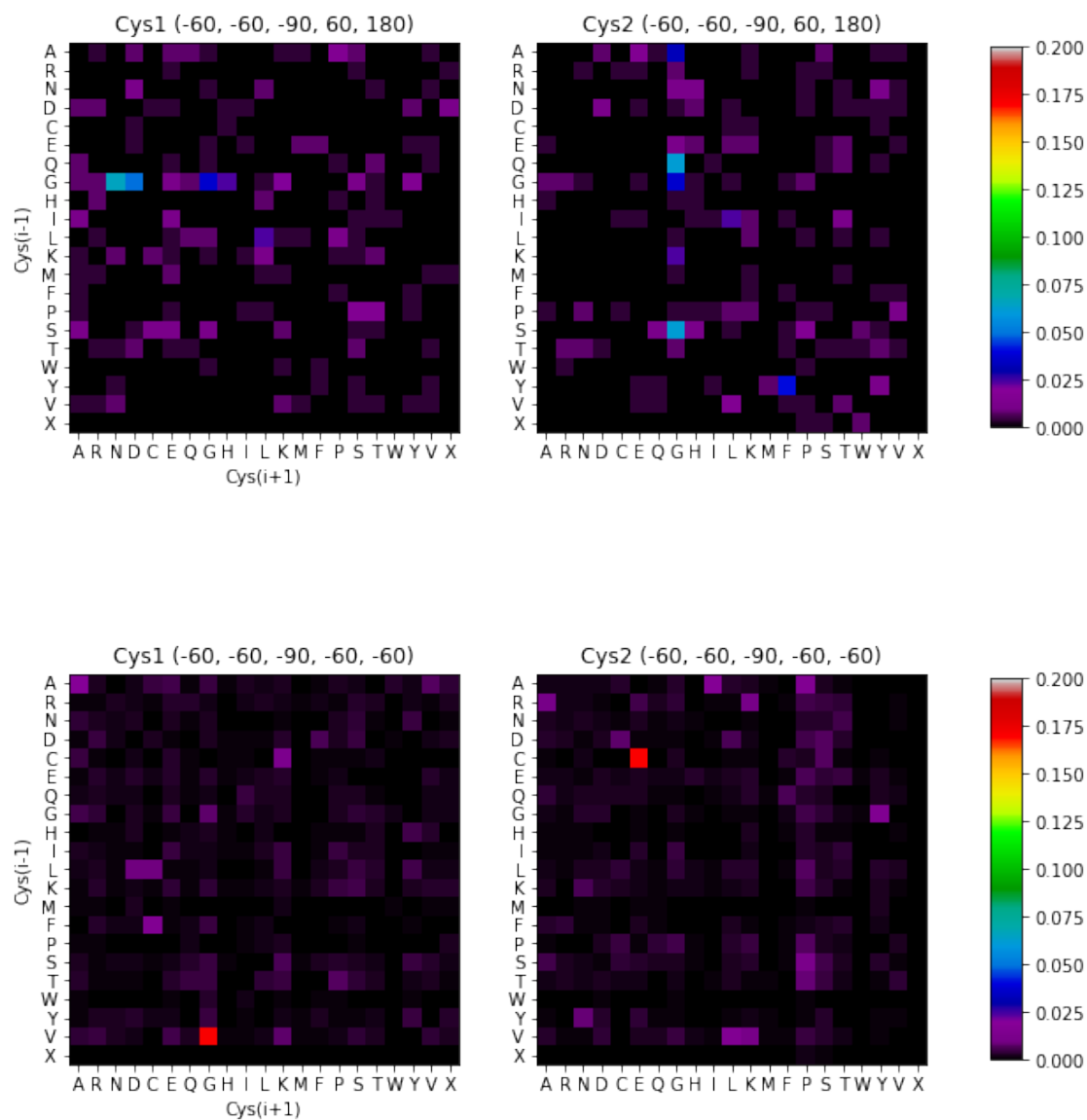


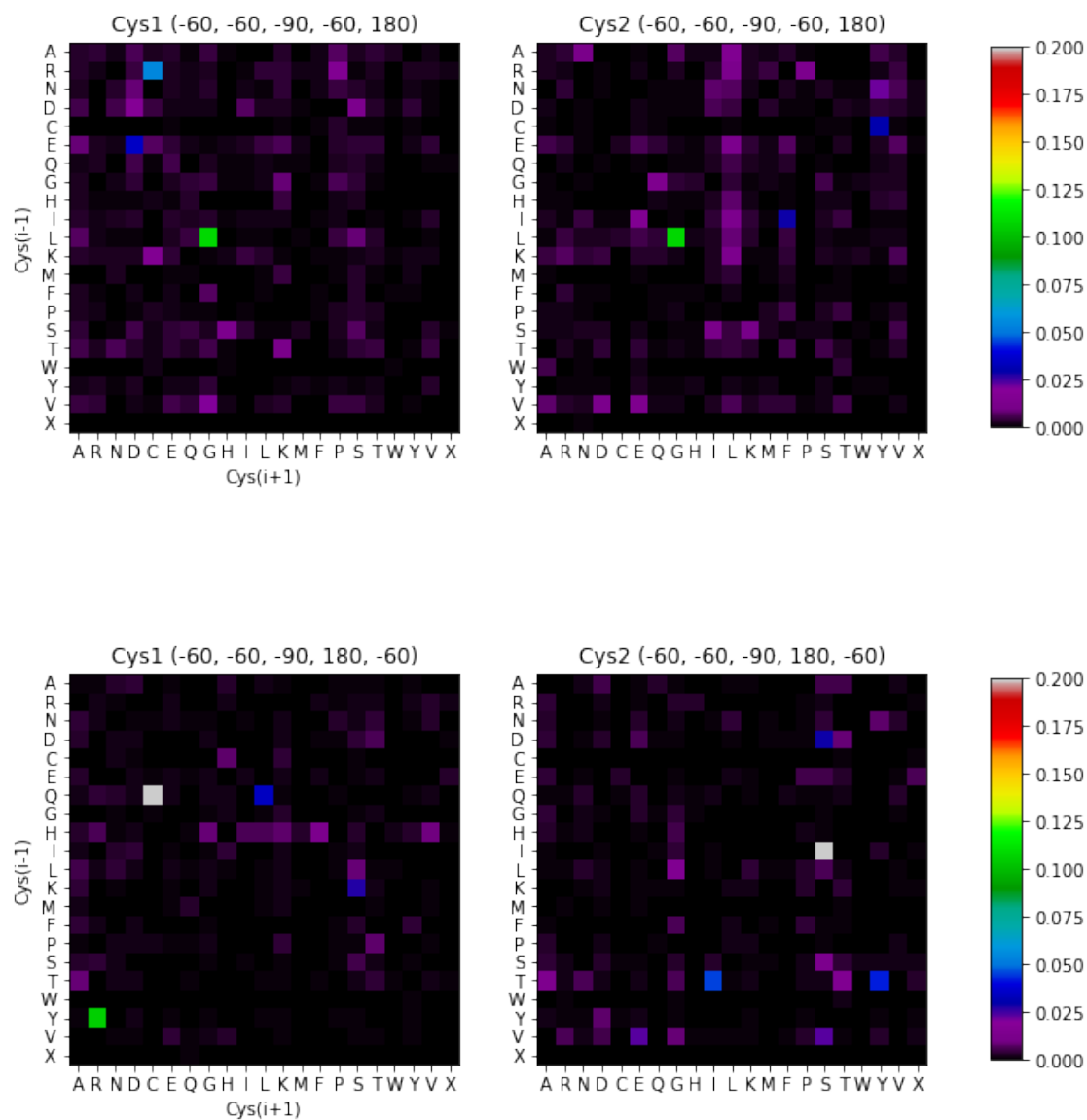


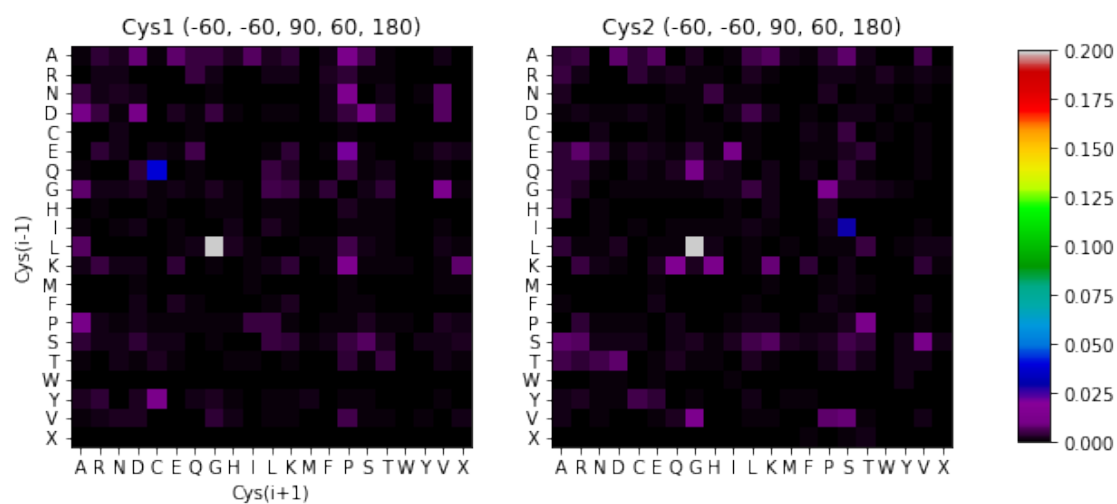
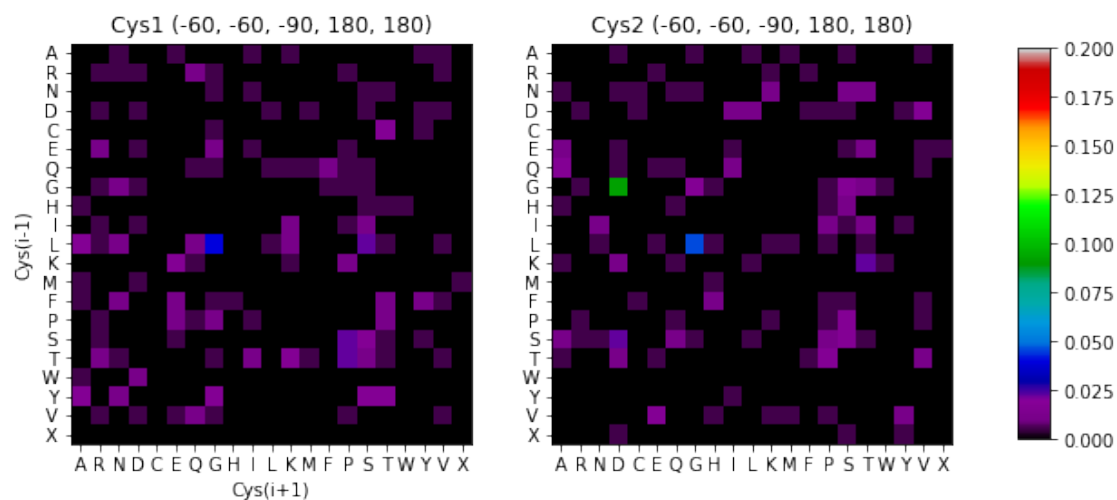


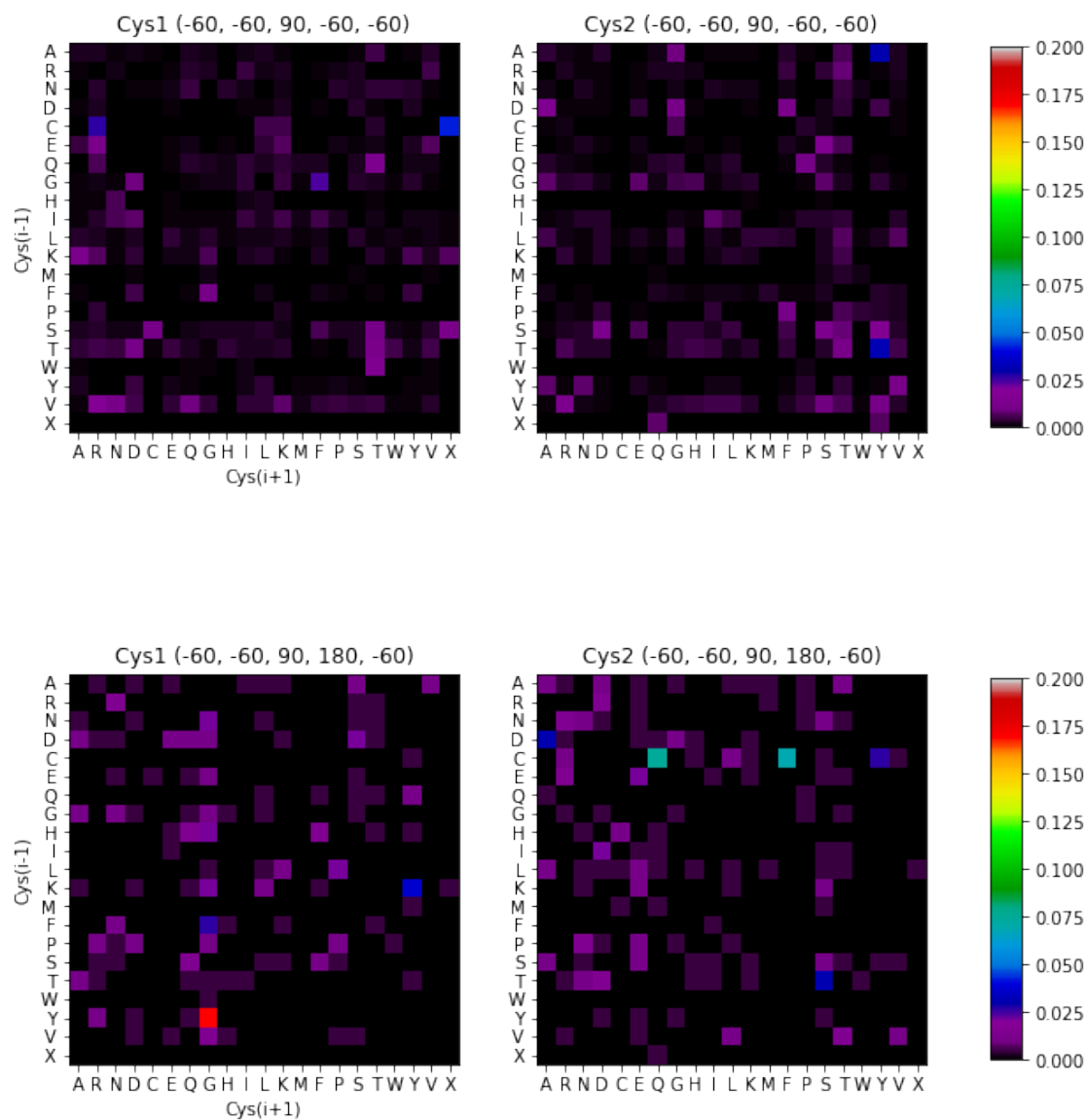


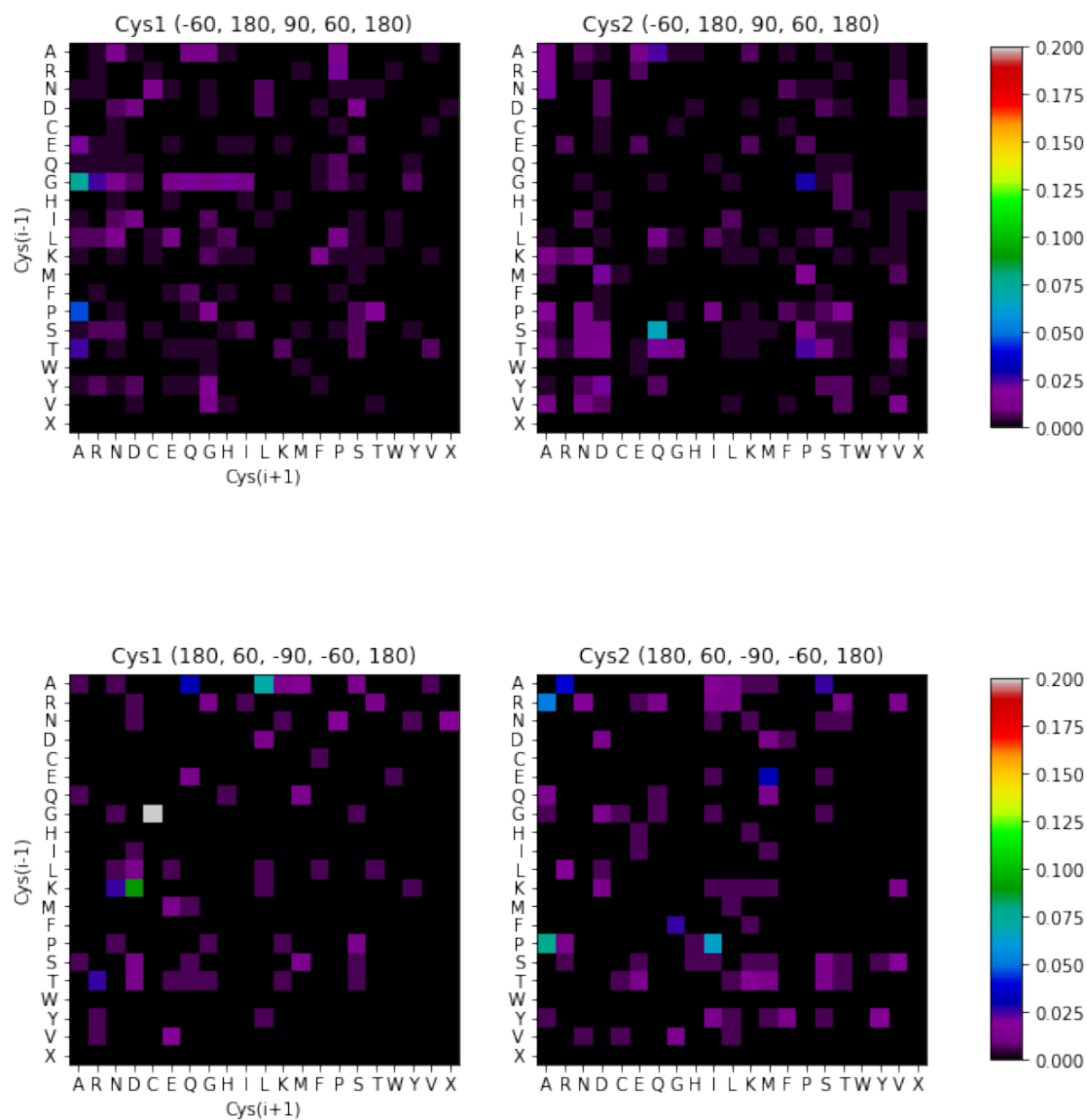


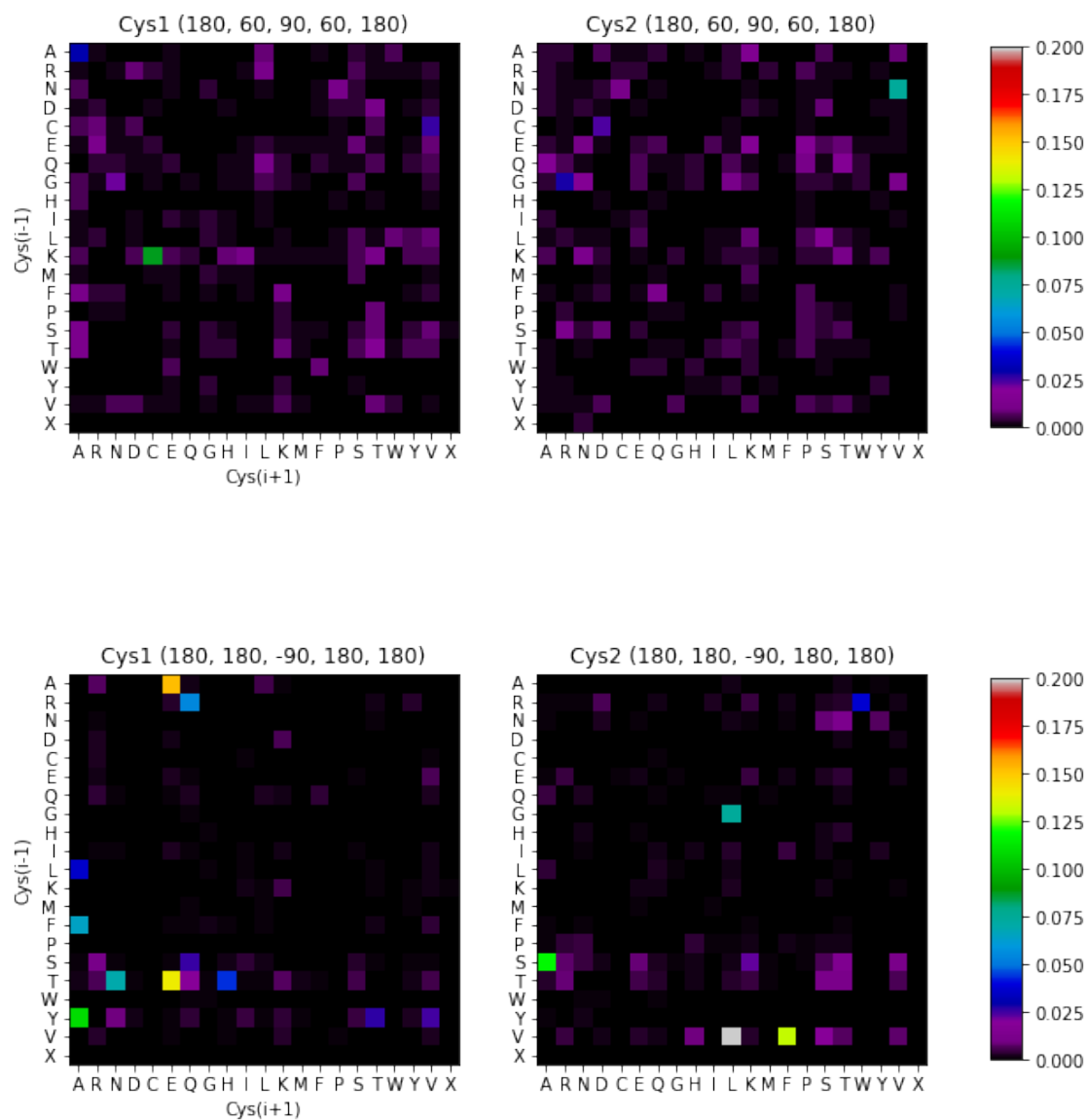




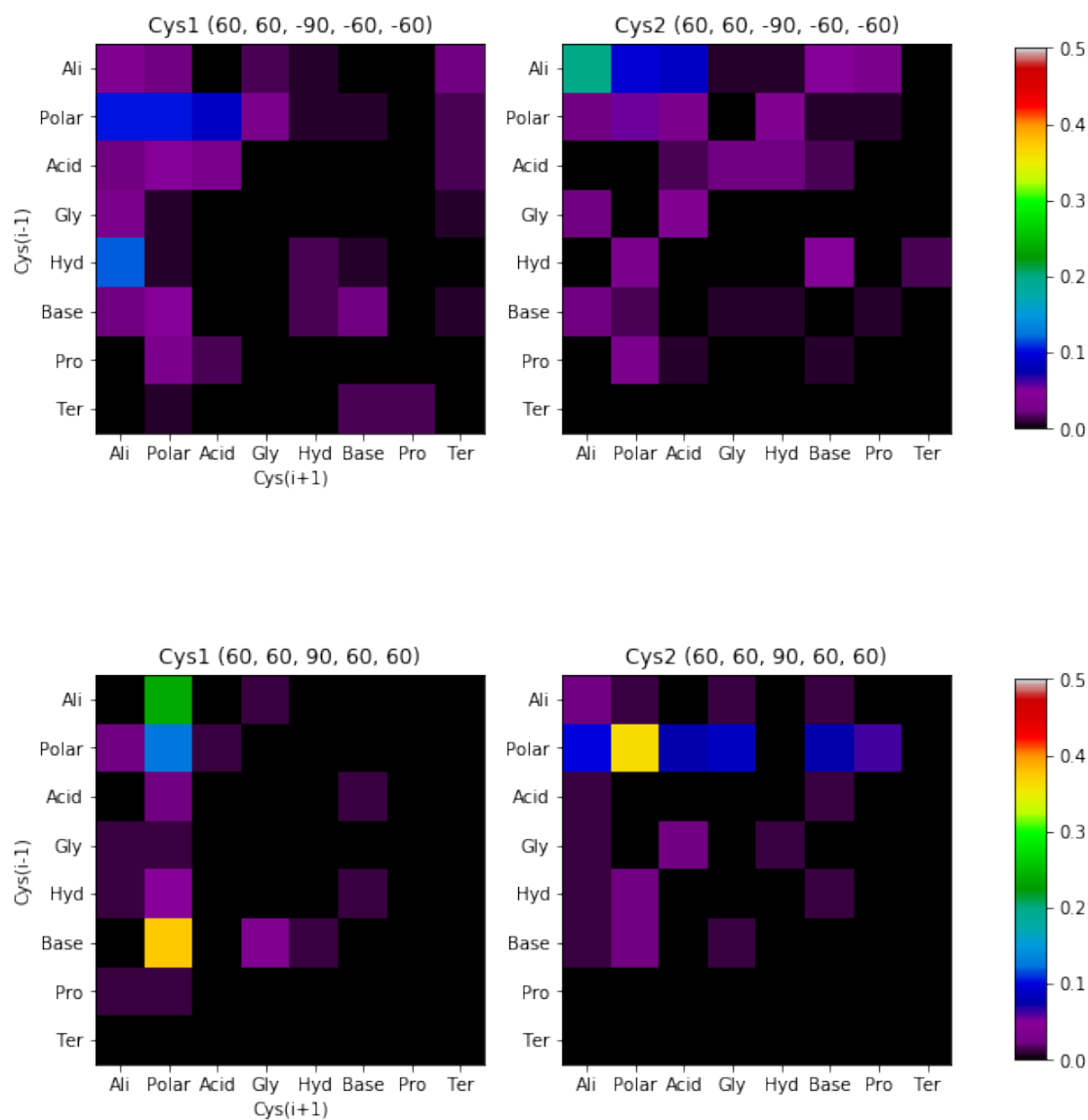




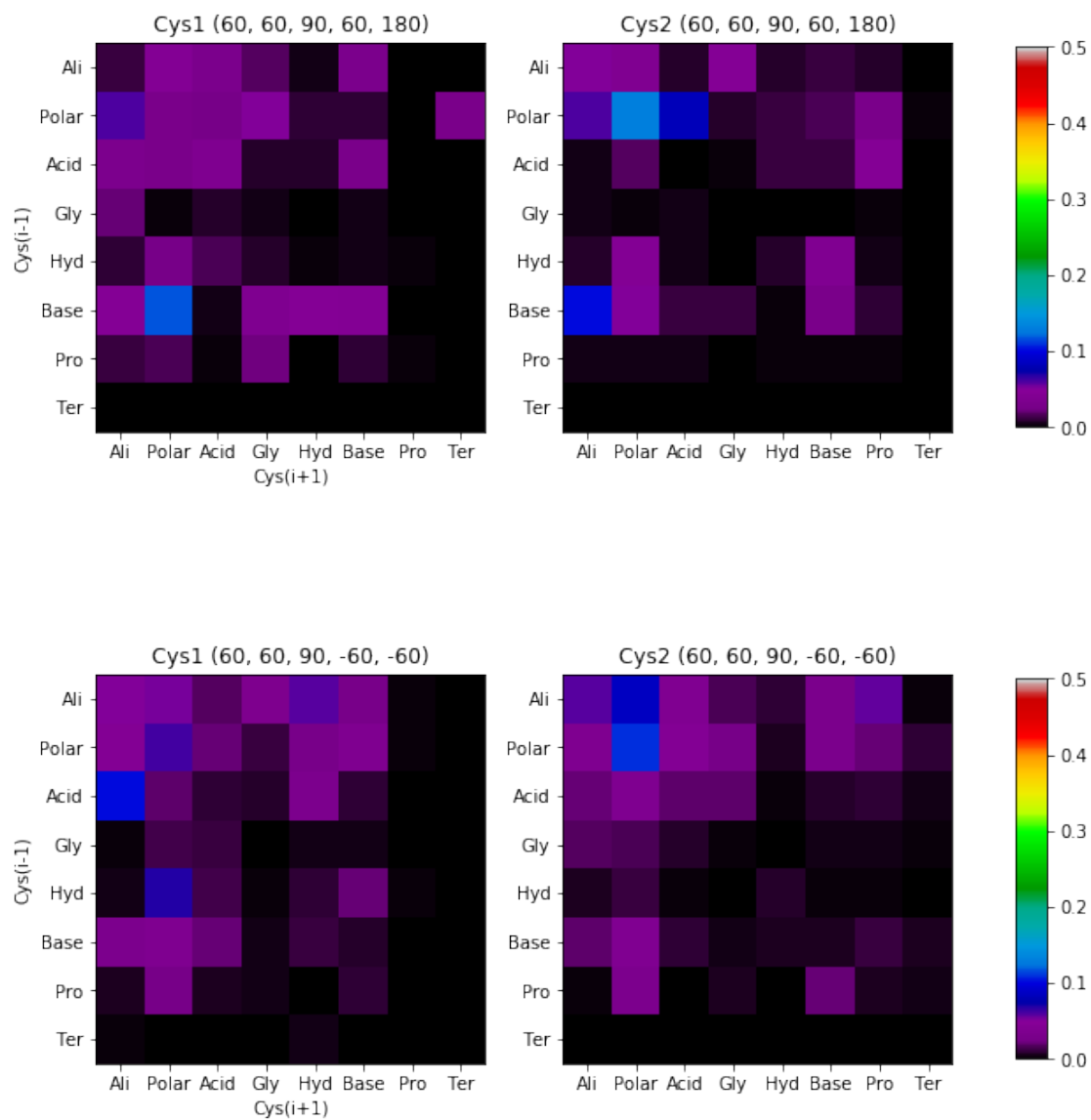


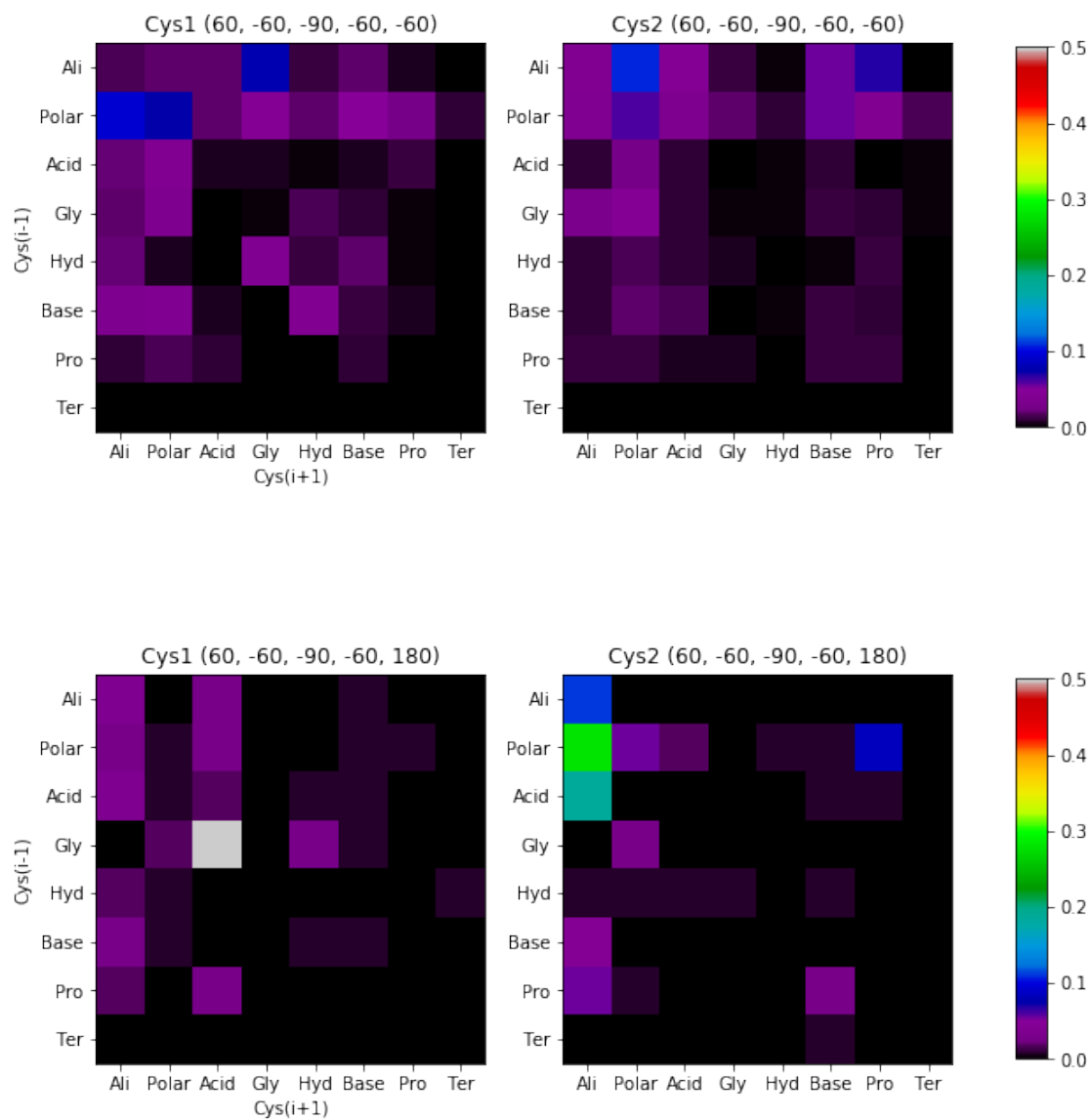


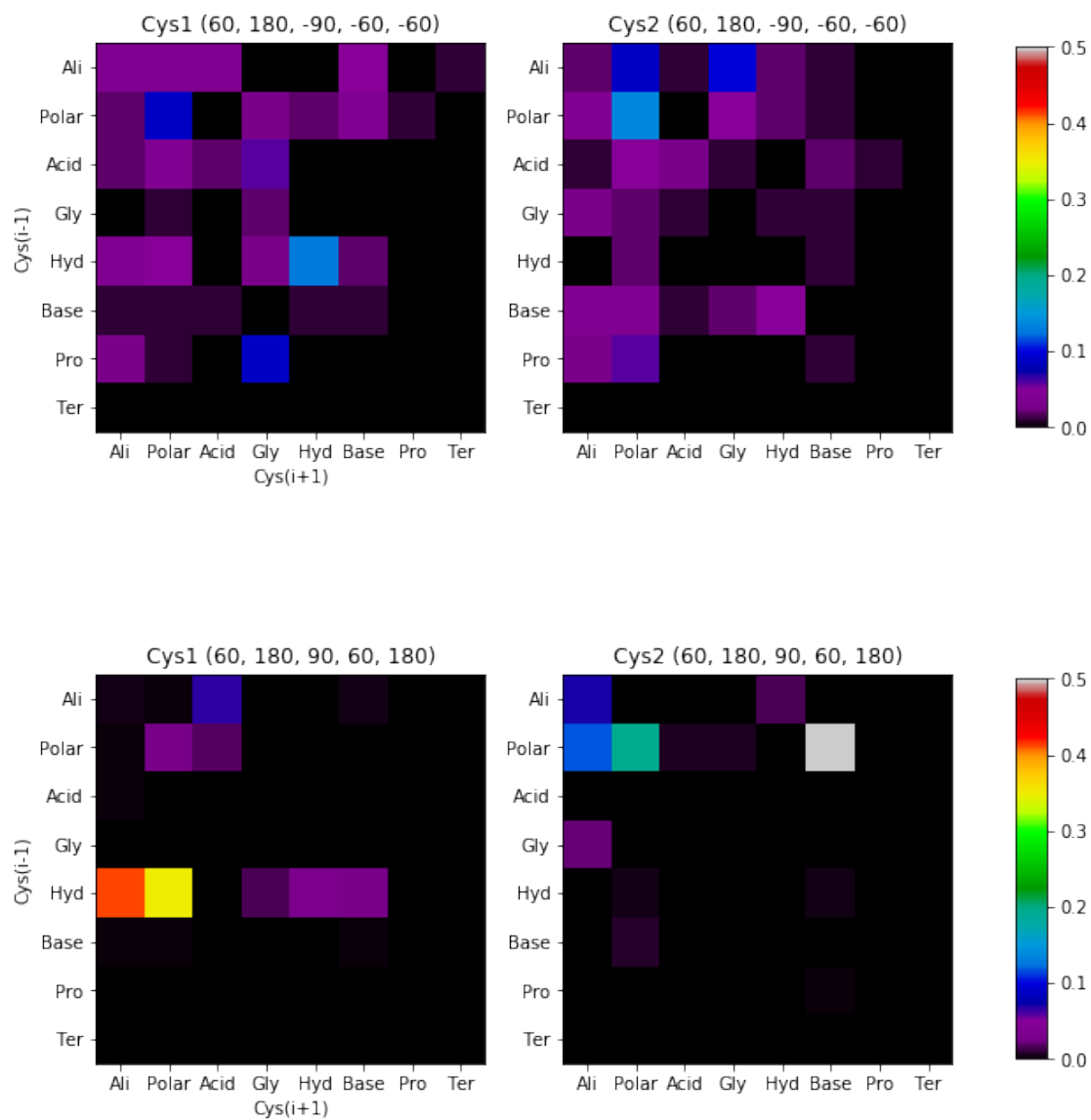
### 1.6.1 Frequency of Amino Acid by Biochemical Property

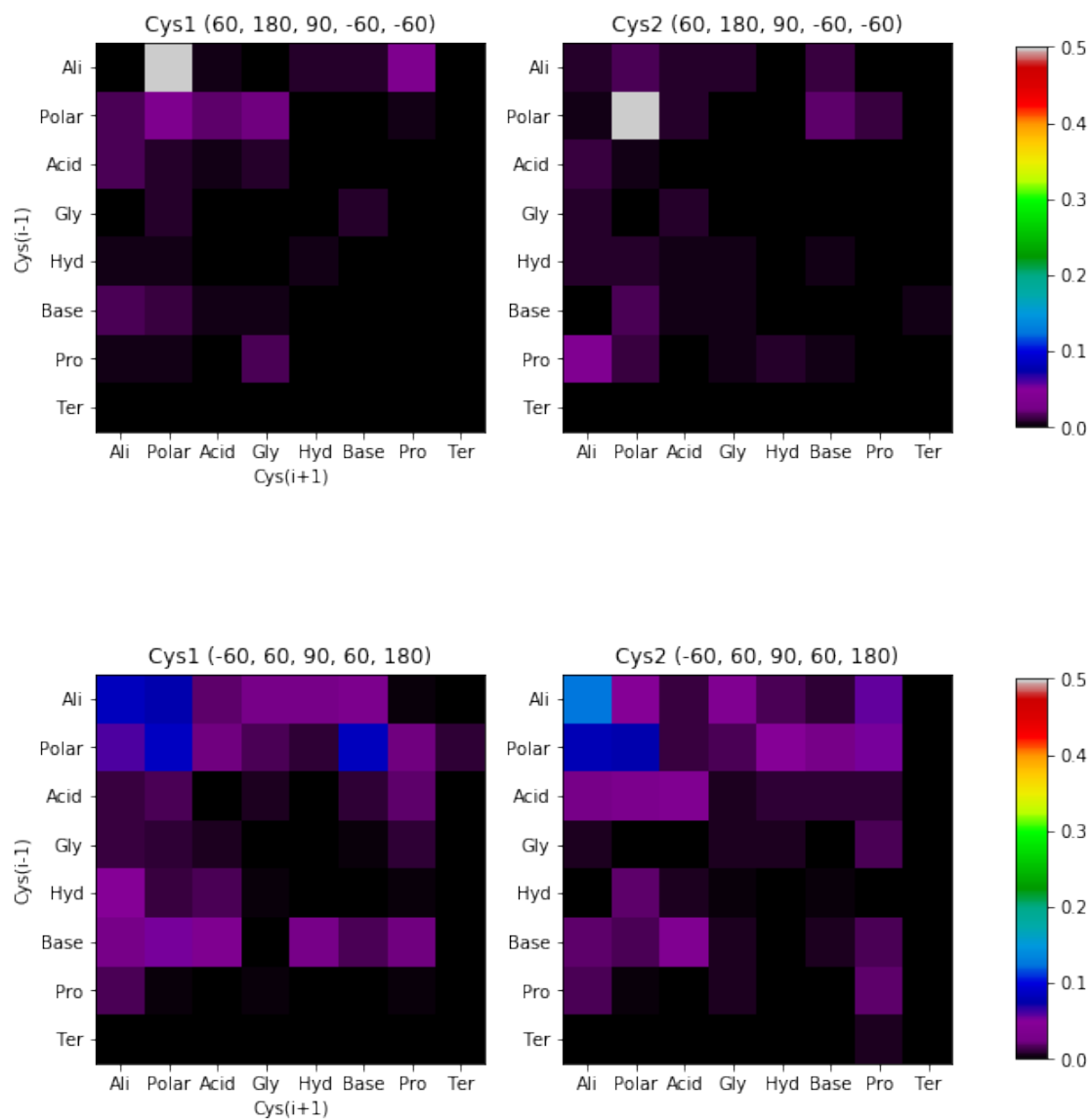


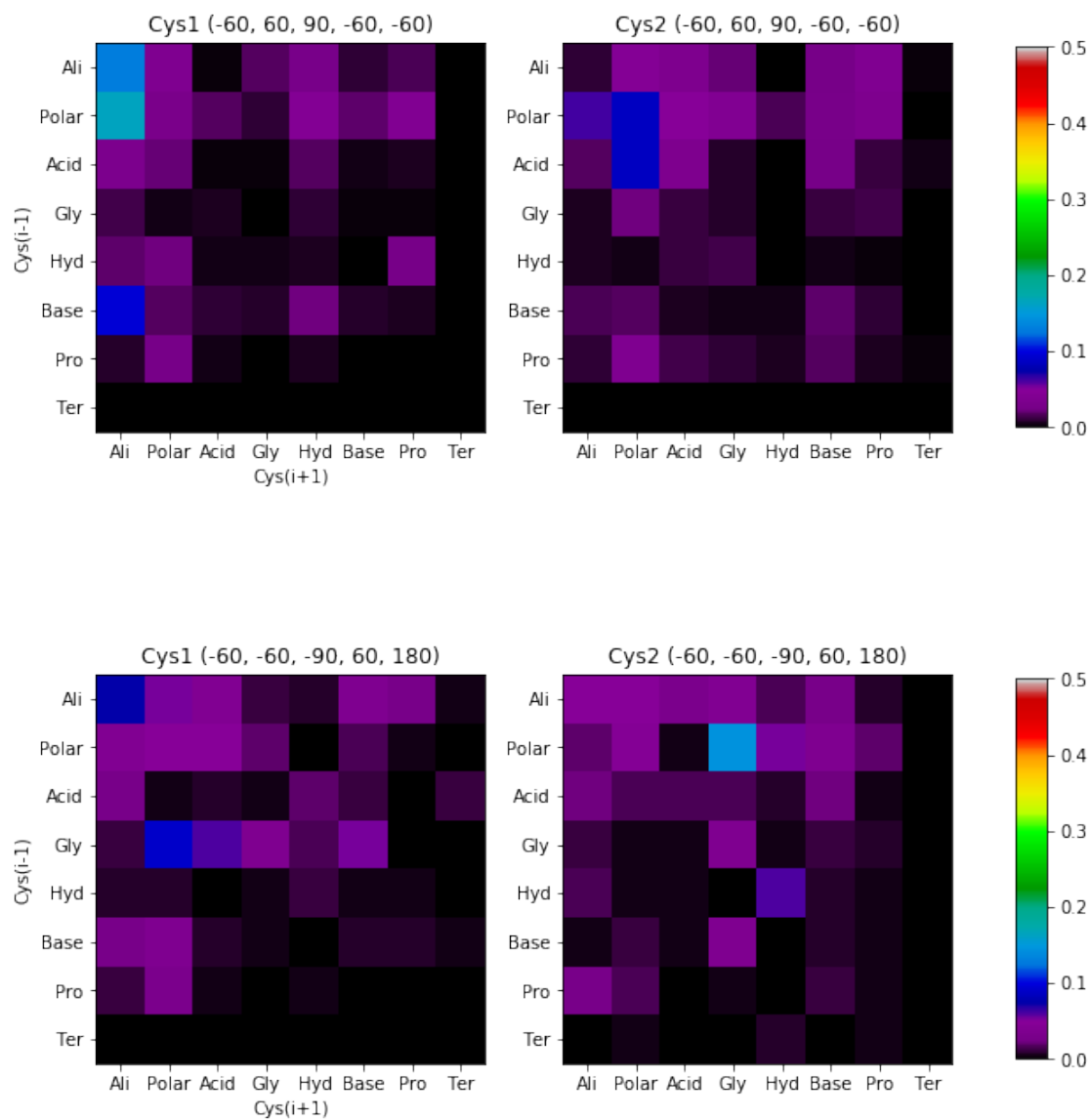


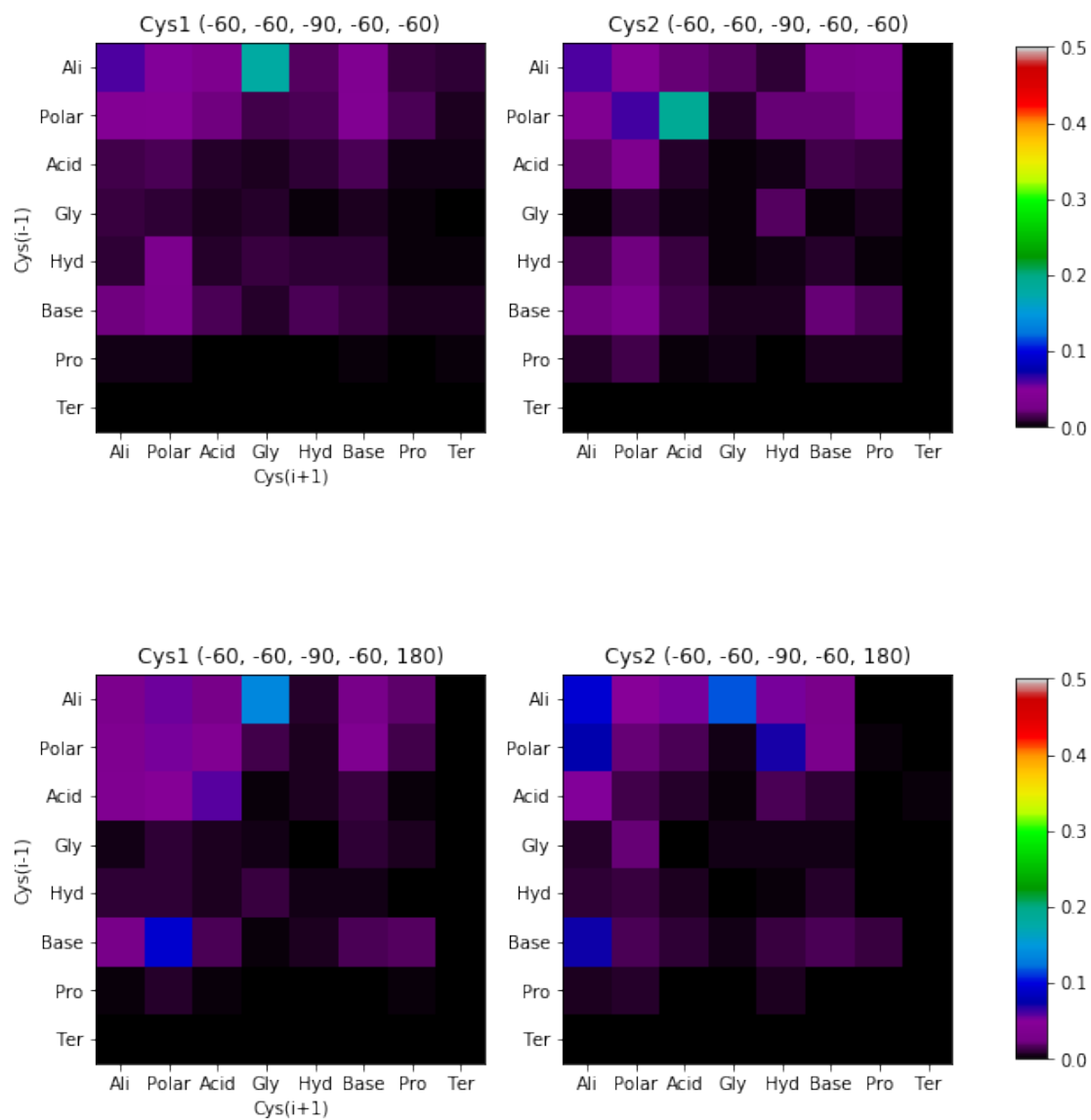


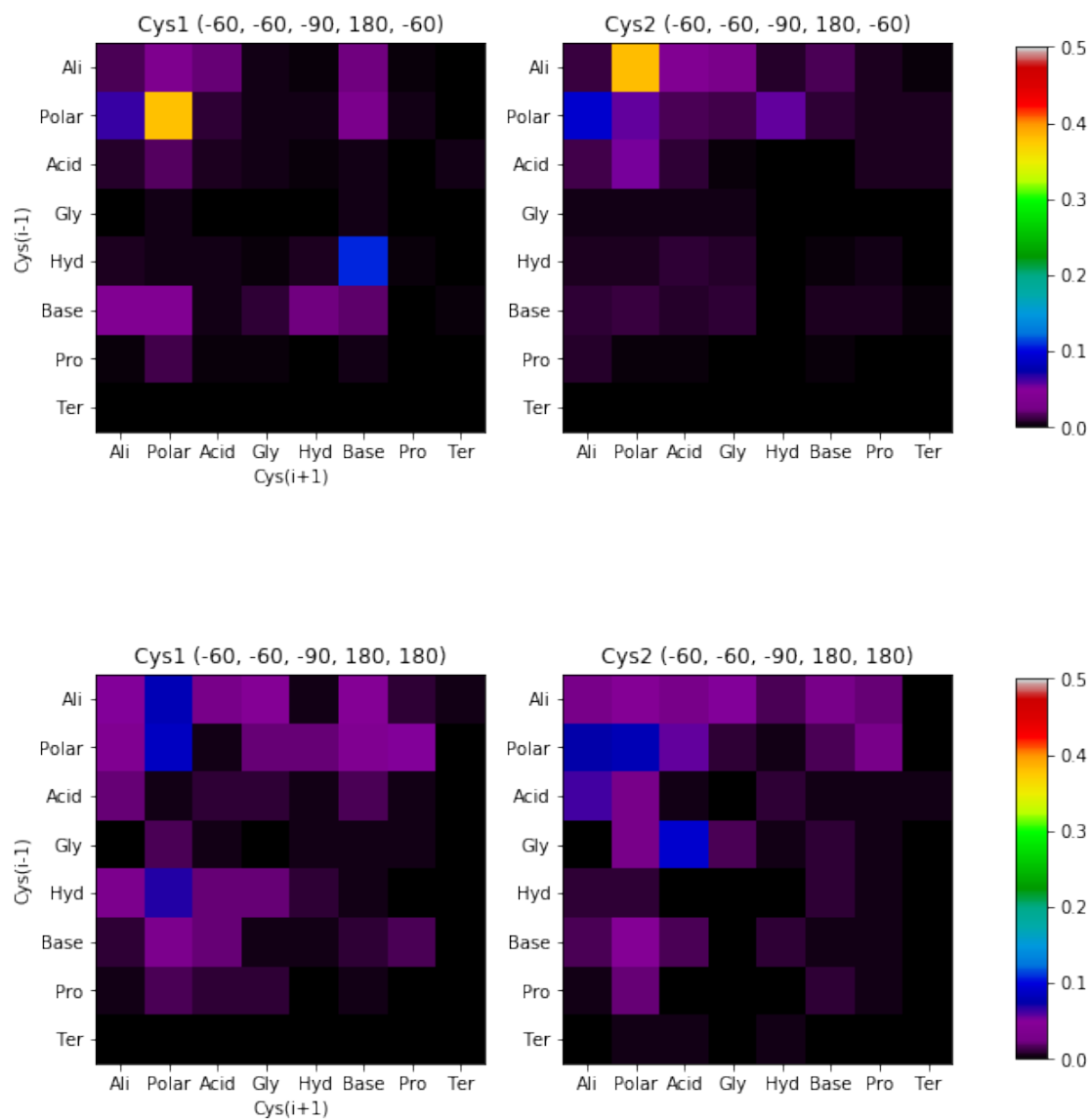


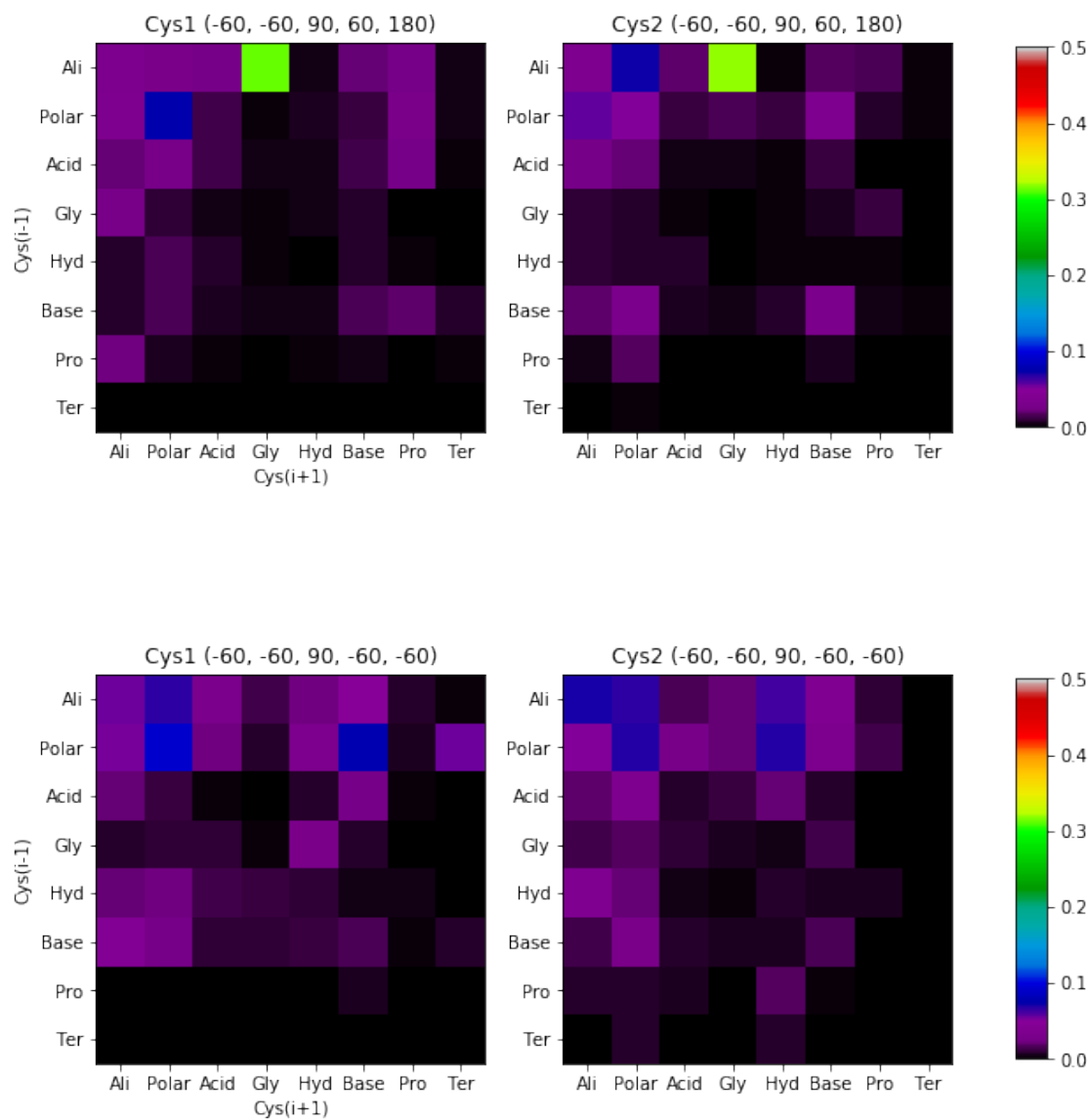




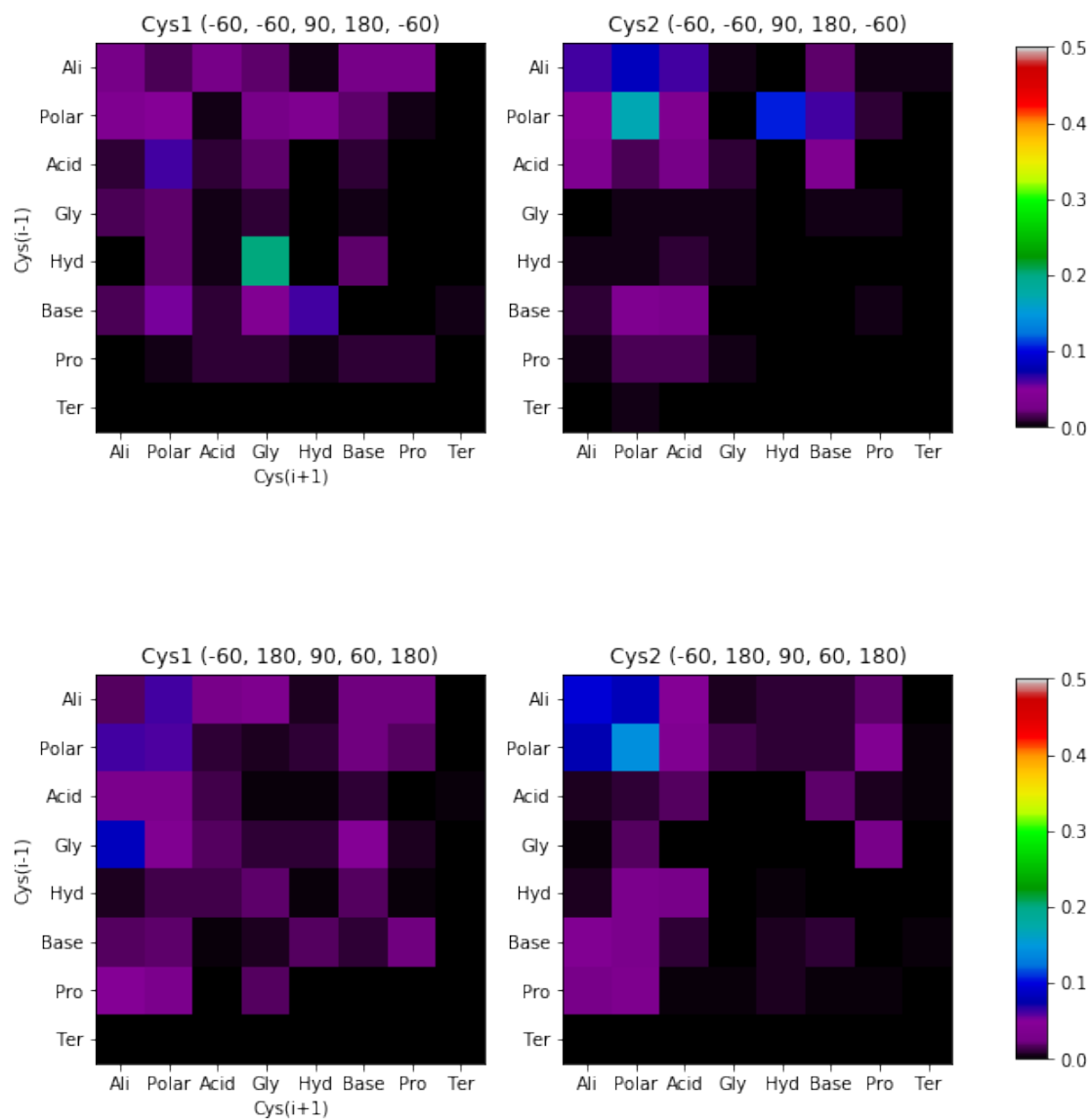


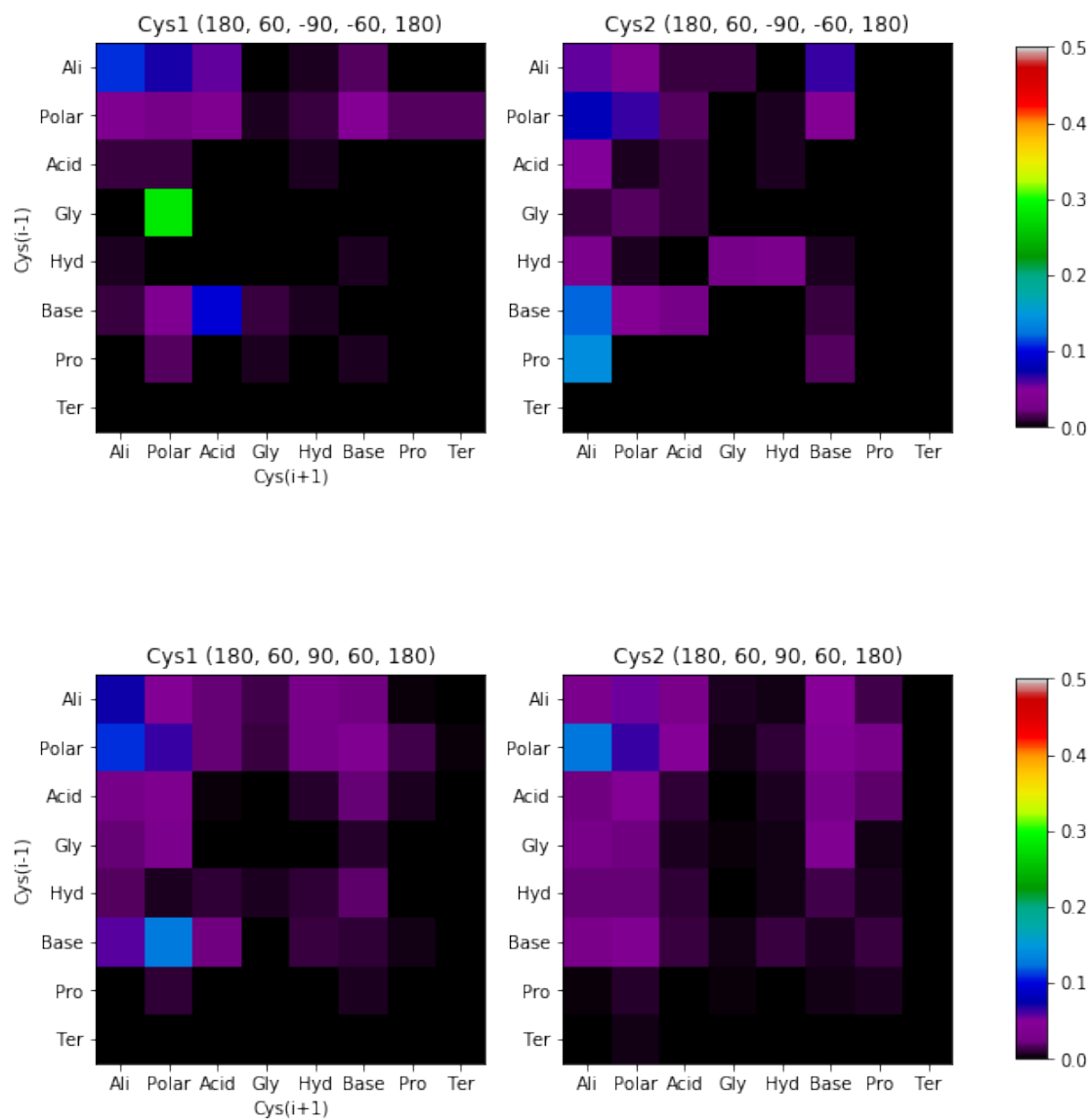


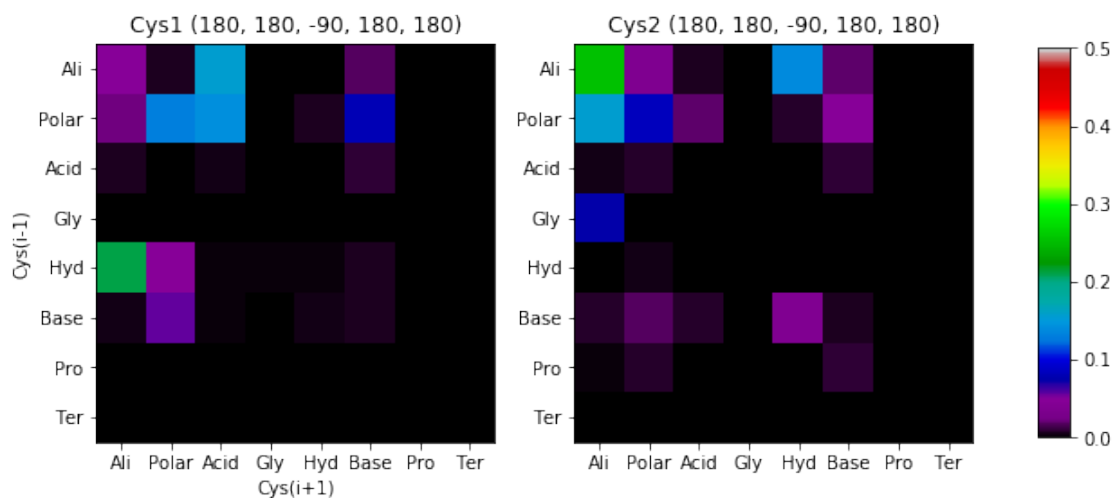




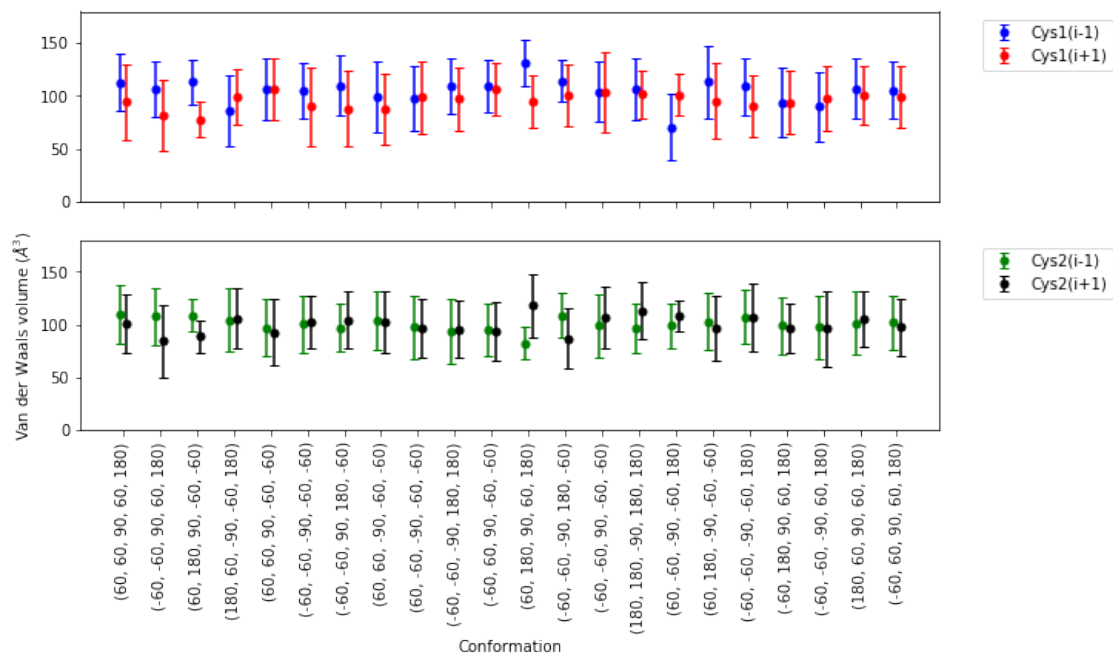








## 1.6.2 Frequency of Amino Acid by VDW Radi



Out[14]: ()

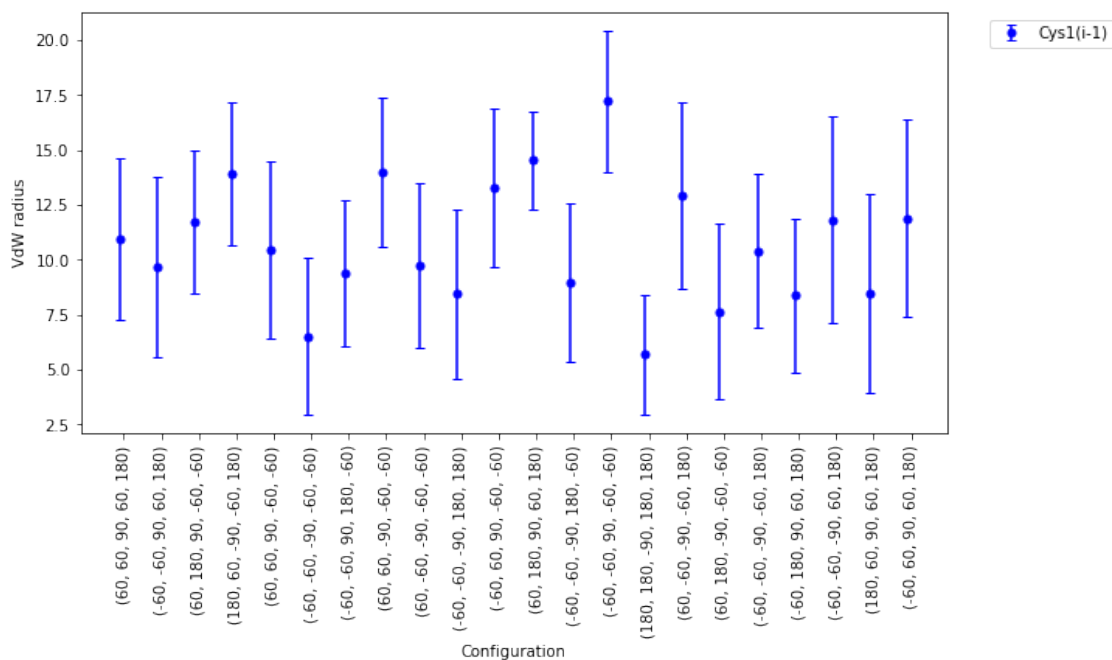
```
(60, 60, -90, -60, -60) 14.020357311082645
(60, 60, 90, 60, 180) 10.94146046003315
(60, 60, 90, -60, -60) 10.43890017018778
(60, -60, -90, -60, -60) 9.729912143632106
```

```

(60, -60, -90, -60, 180) 12.91536445628846
(60, 180, -90, -60, -60) 7.642602000650002
(60, 180, 90, 60, 180) 14.546032212688422
(60, 180, 90, -60, -60) 11.707496947589744
(-60, 60, 90, 60, 180) 11.888714858685315
(-60, 60, 90, -60, -60) 13.290180158957144
(-60, -60, -90, 60, 180) 11.819749037119658
(-60, -60, -90, -60, -60) 6.524539662137568
(-60, -60, -90, -60, 180) 10.409521748311489
(-60, -60, -90, 180, -60) 8.957282444078768
(-60, -60, -90, 180, 180) 8.447279527899408
(-60, -60, 90, 60, 180) 9.673237960921325
(-60, -60, 90, -60, -60) 17.22122359748012
(-60, -60, 90, 180, -60) 9.405272451410811
(-60, 180, 90, 60, 180) 8.379706433324529
(180, 60, -90, -60, 180) 13.951319572960529
(180, 60, 90, 60, 180) 8.473521595474157
(180, 180, -90, 180, 180) 5.687167499307144

```

Out[15]: ( )



```

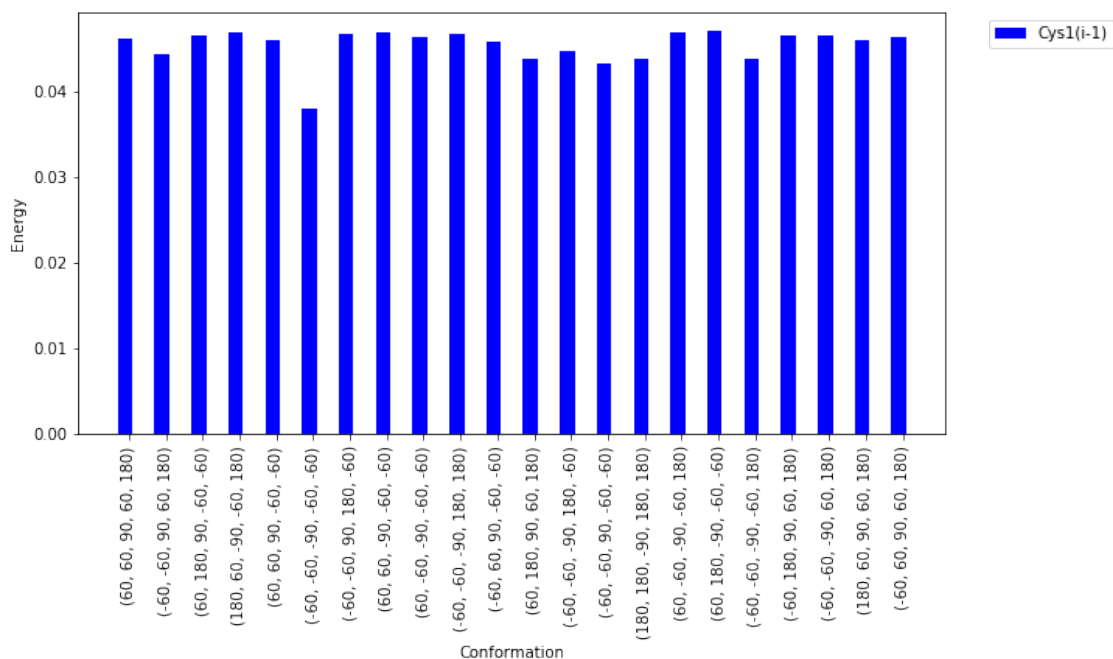
(60, 60, -90, -60, -60) 0.0469502309429
(60, 60, 90, 60, 180) 0.0462157352289
(60, 60, 90, -60, -60) 0.0459843158415
(60, -60, -90, -60, -60) 0.0464094741215

```

```

(60, -60, -90, -60, 180) 0.0470046519656
(60, 180, -90, -60, -60) 0.0470228063209
(60, 180, 90, 60, 180) 0.0438089678949
(60, 180, 90, -60, -60) 0.0465650498305
(-60, 60, 90, 60, 180) 0.0464363634417
(-60, 60, 90, -60, -60) 0.045839530003
(-60, -60, -90, 60, 180) 0.0465980268482
(-60, -60, -90, -60, -60) 0.038086471394
(-60, -60, -90, -60, 180) 0.0438315294846
(-60, -60, -90, 180, -60) 0.0446946581
(-60, -60, -90, 180, 180) 0.0468054160543
(-60, -60, 90, 60, 180) 0.0444422438124
(-60, -60, 90, -60, -60) 0.0433769137097
(-60, -60, 90, 180, -60) 0.0467602530691
(-60, 180, 90, 60, 180) 0.0464842051513
(180, 60, -90, -60, 180) 0.046859669219
(180, 60, 90, 60, 180) 0.0459665624221
(180, 180, -90, 180, 180) 0.043856925162

```



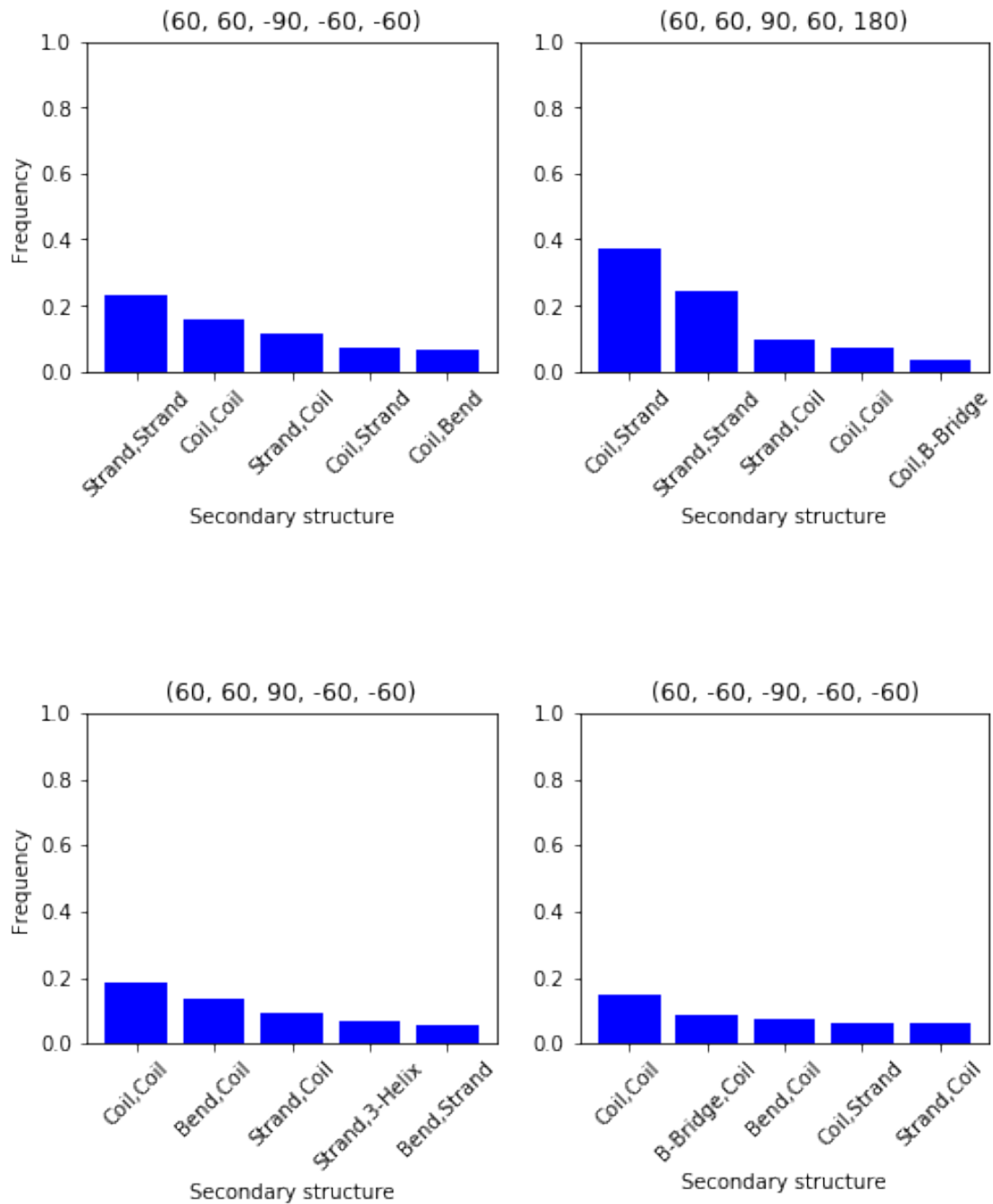
Out[16]: ()

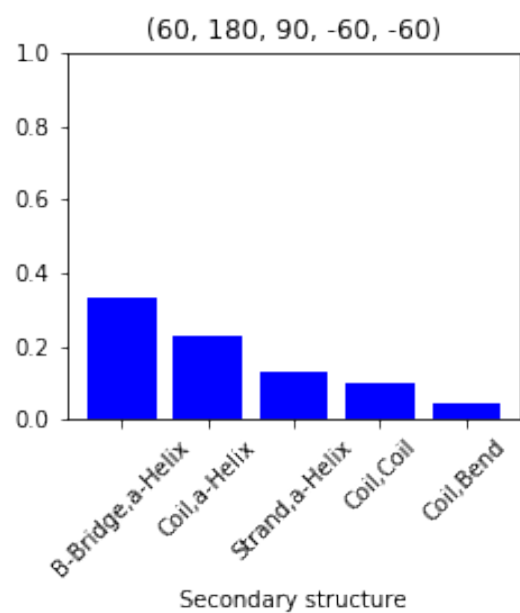
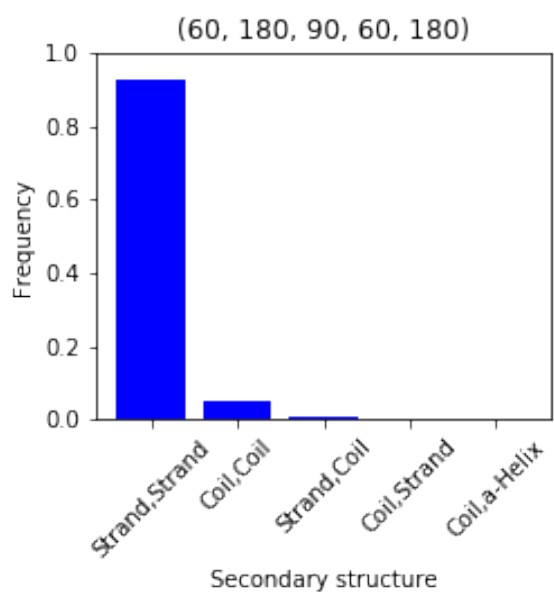
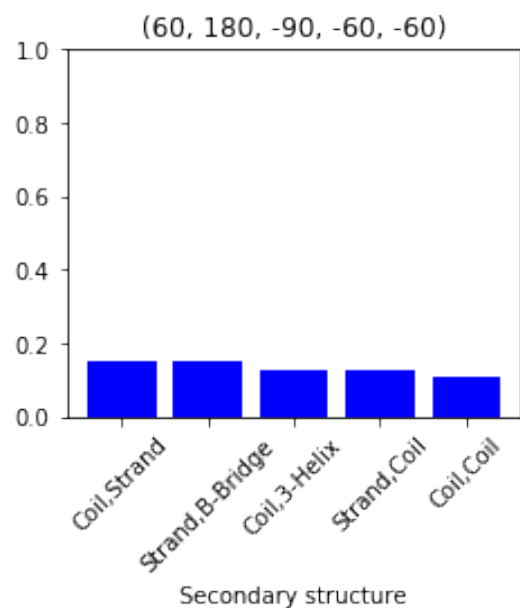
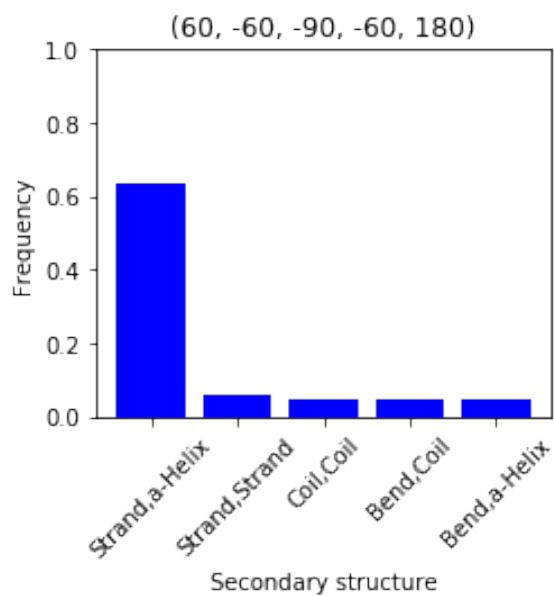
## 1.7 Secondary Structure

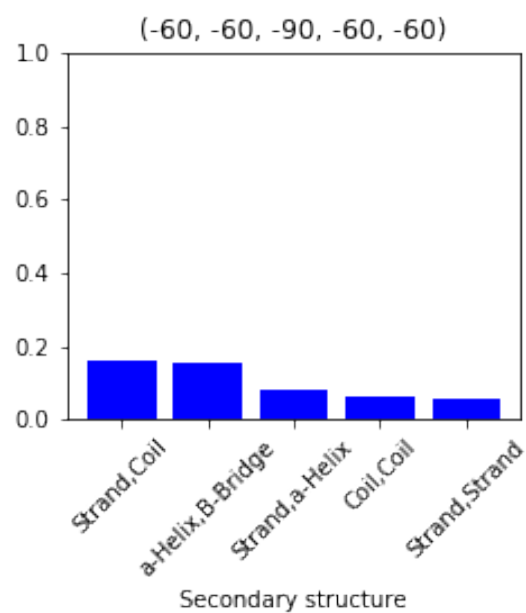
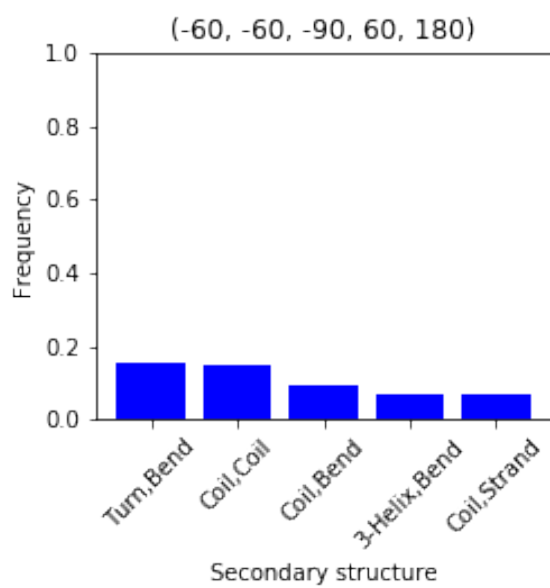
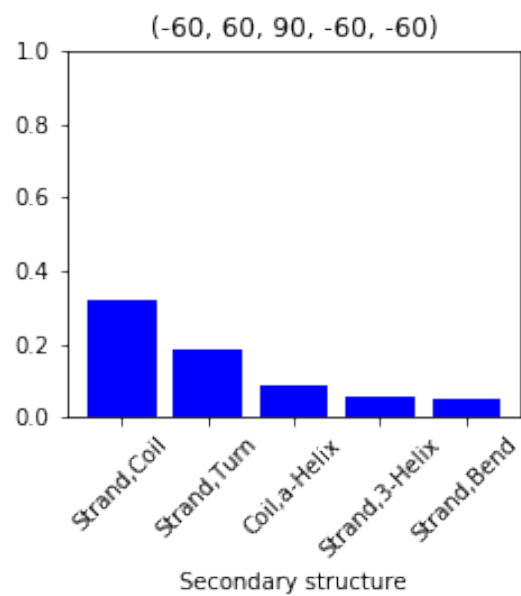
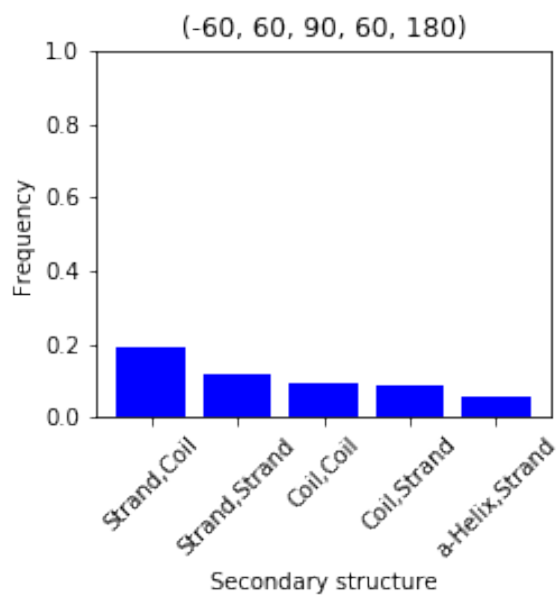
/home/david/coding/lib/python2.7/site-packages/ipykernel\_launcher.py:90: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

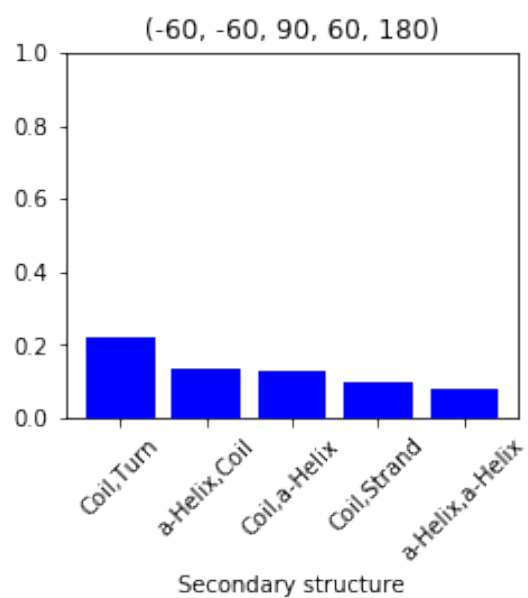
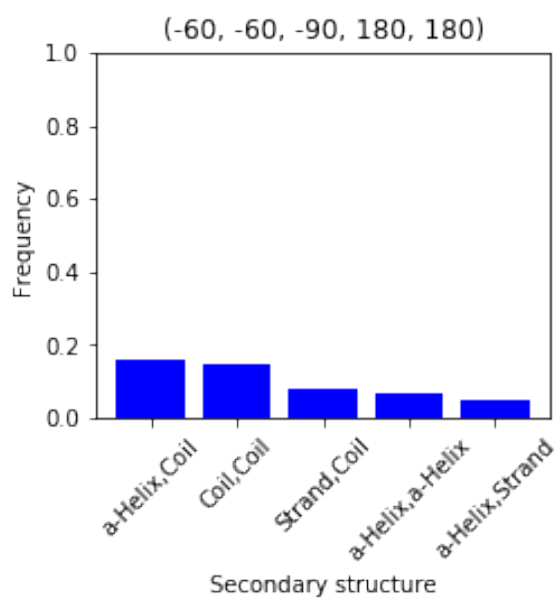
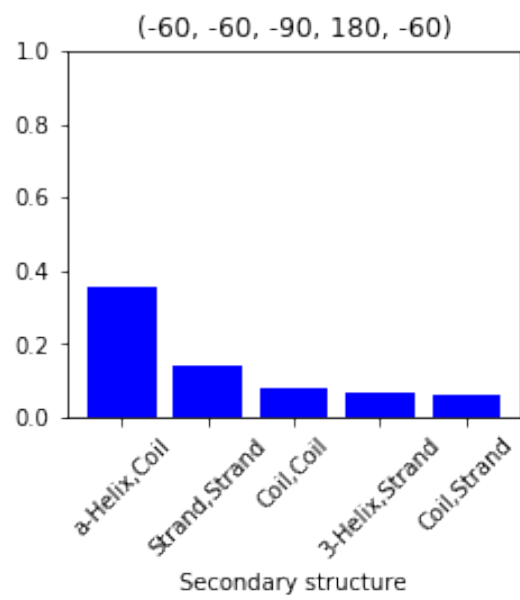
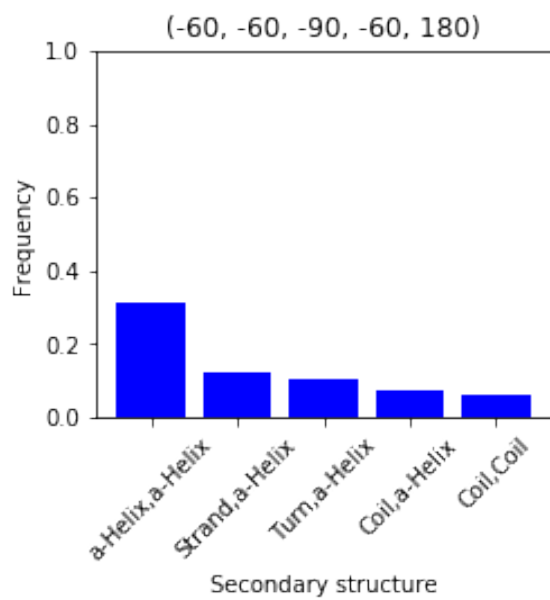
See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#>

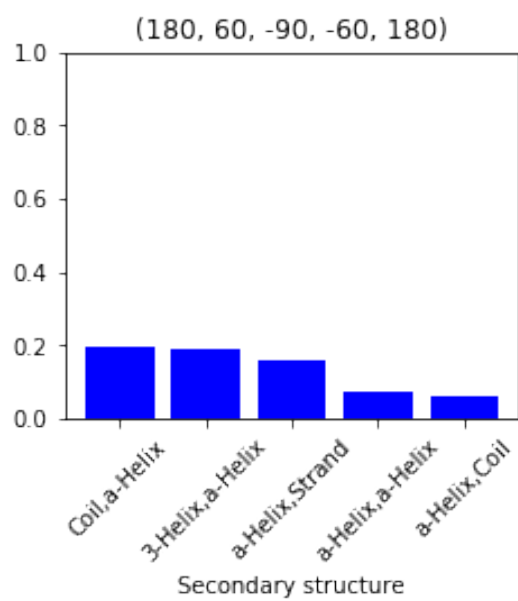
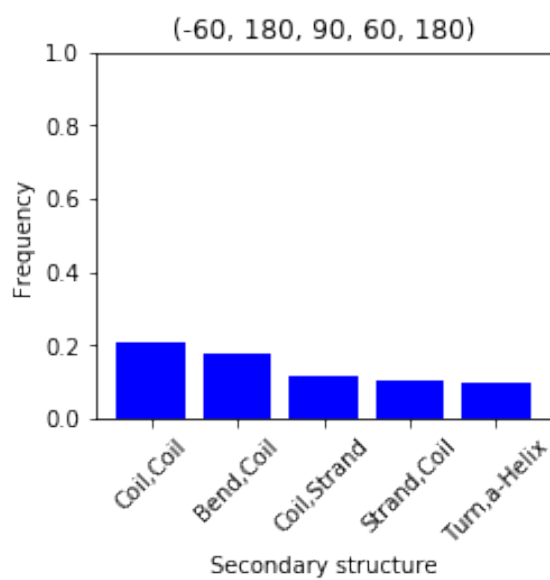
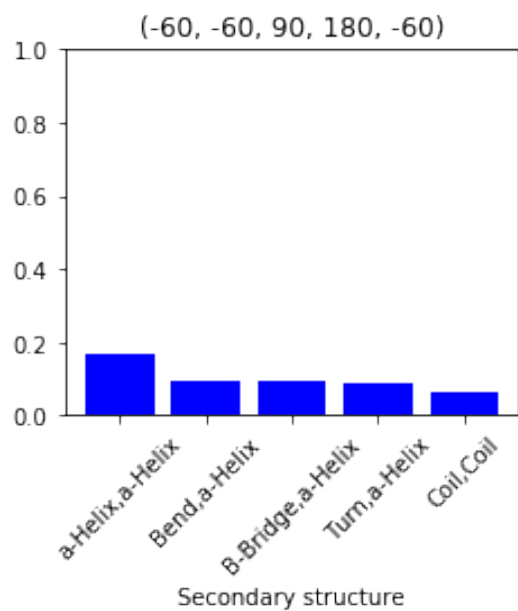
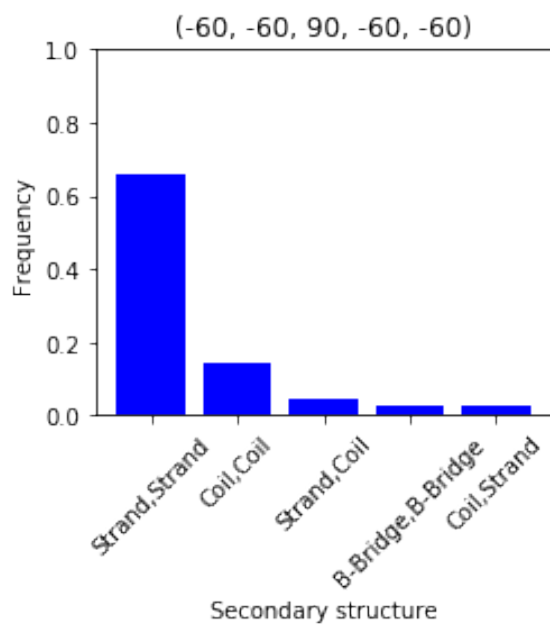


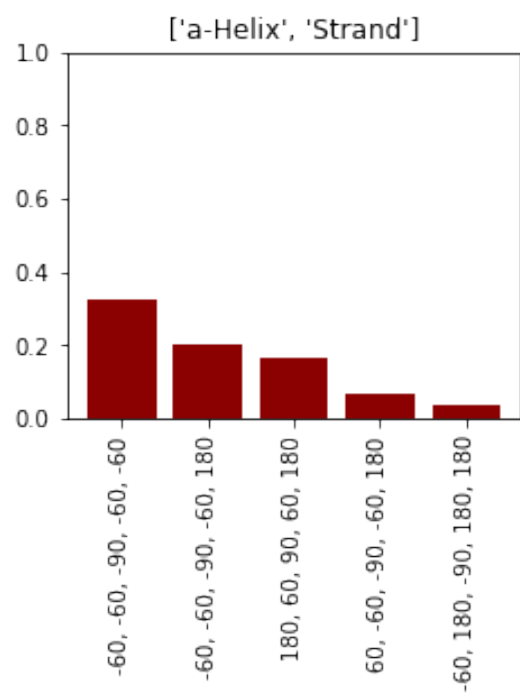
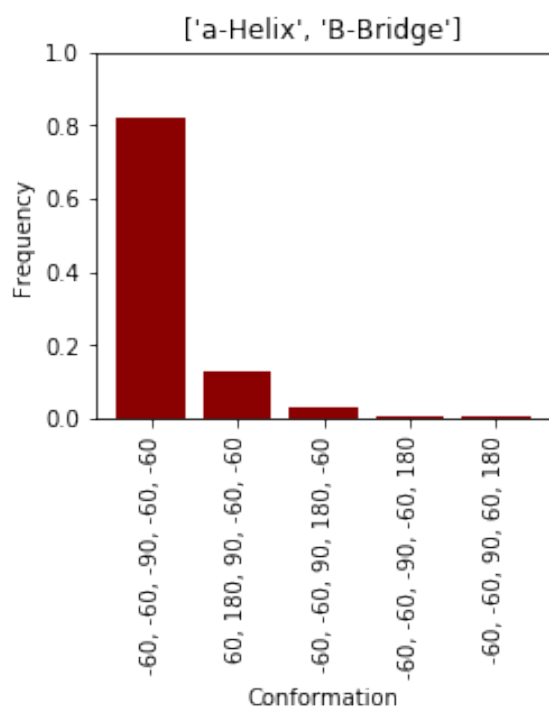
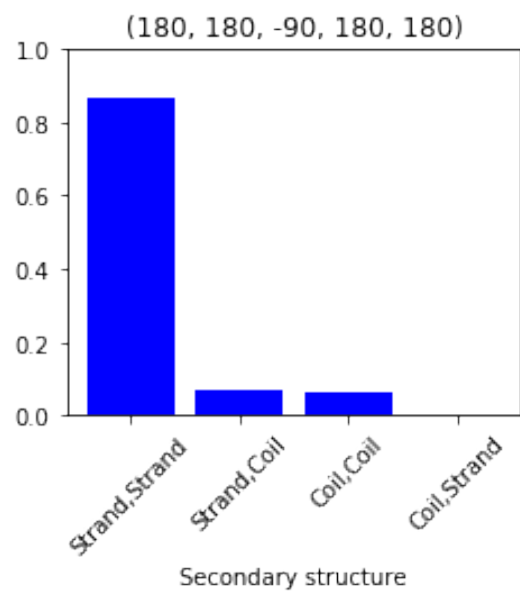
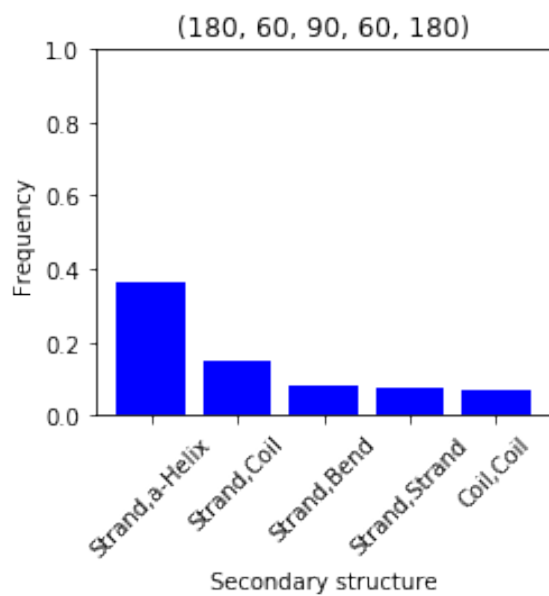


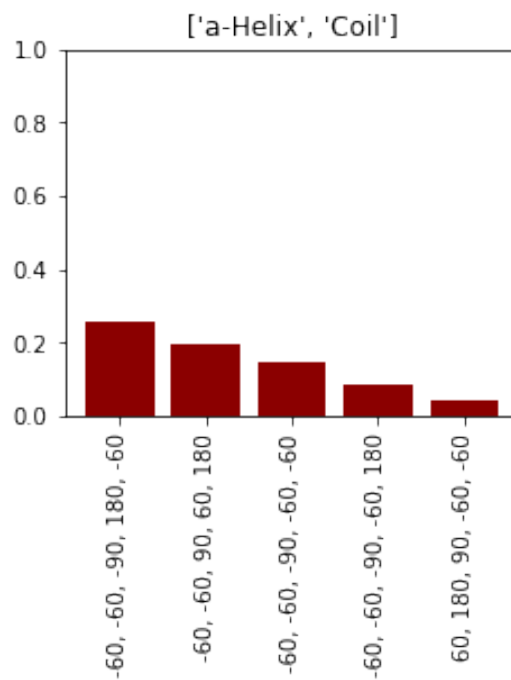
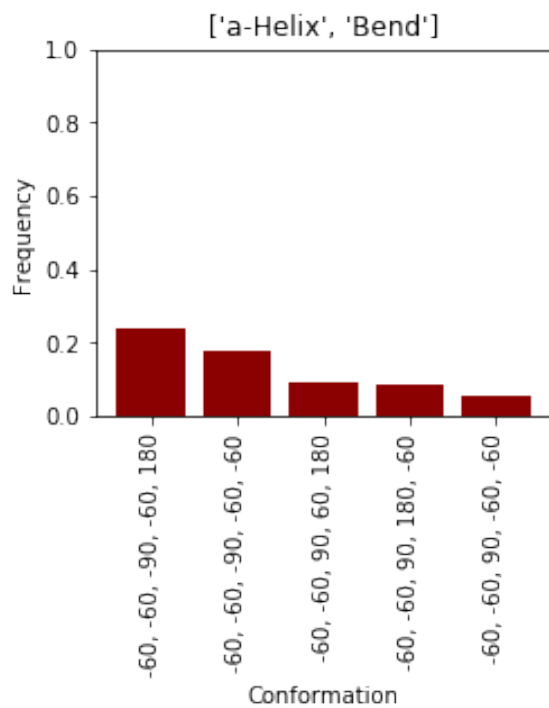
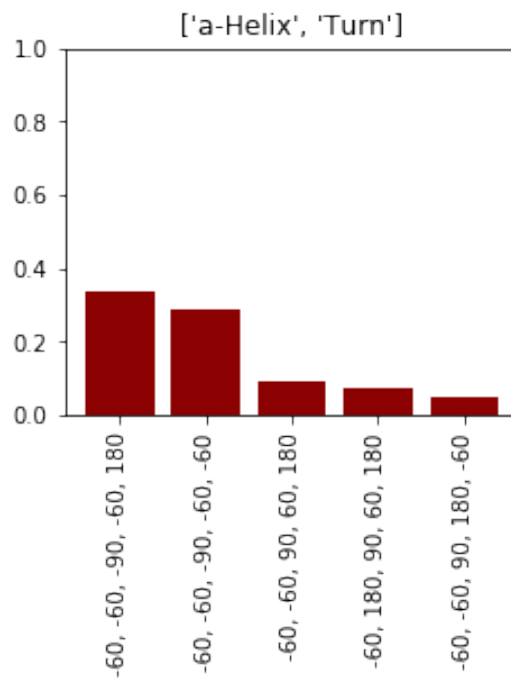
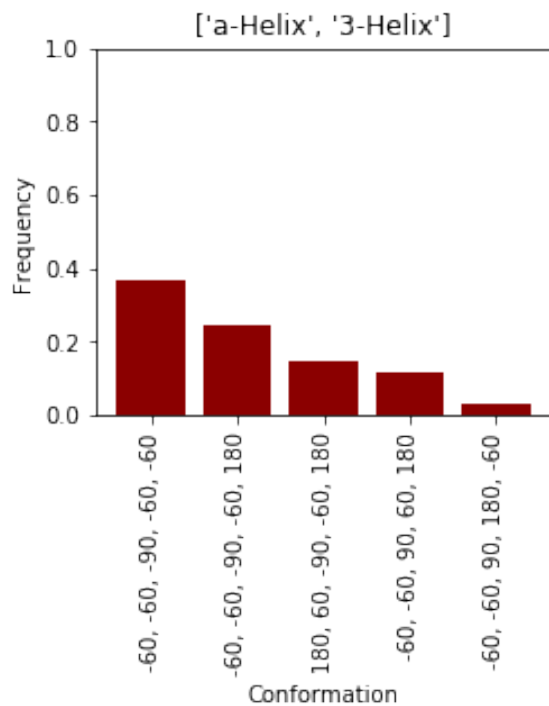


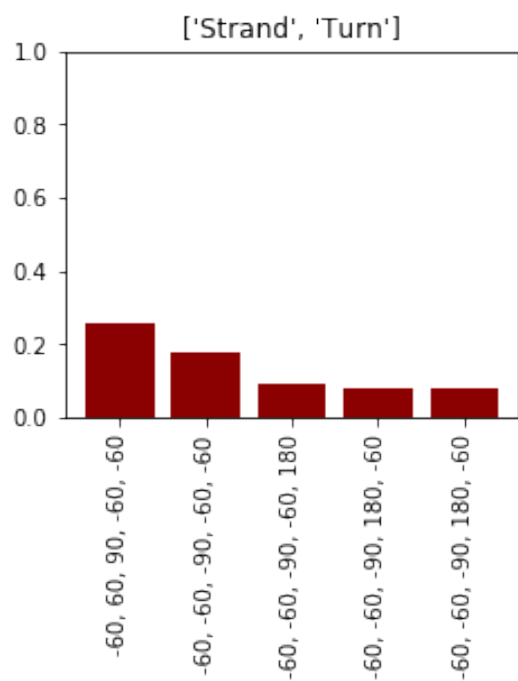
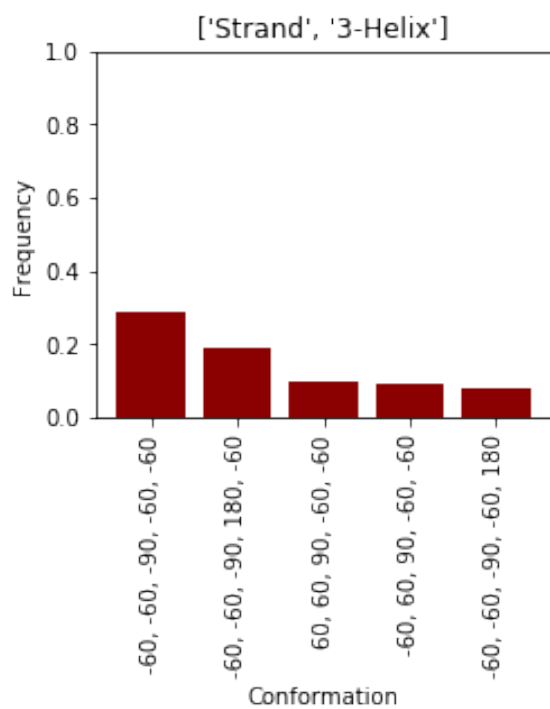
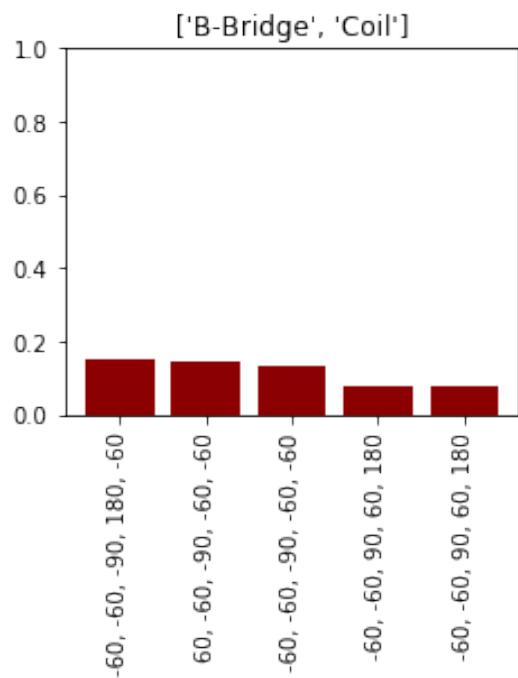
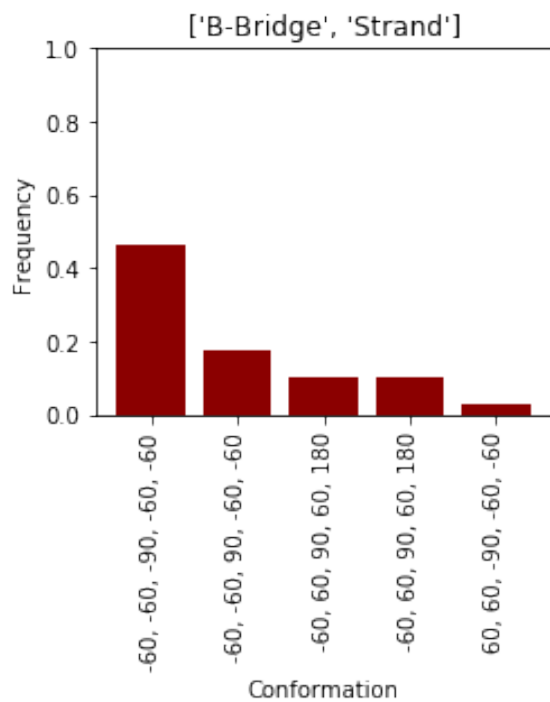


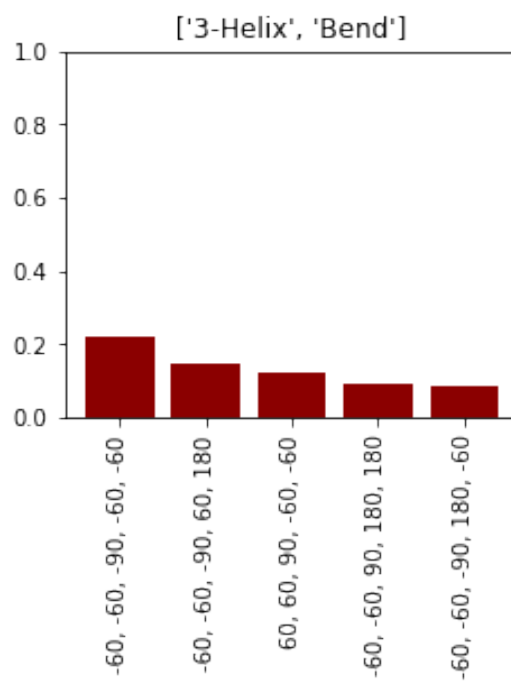
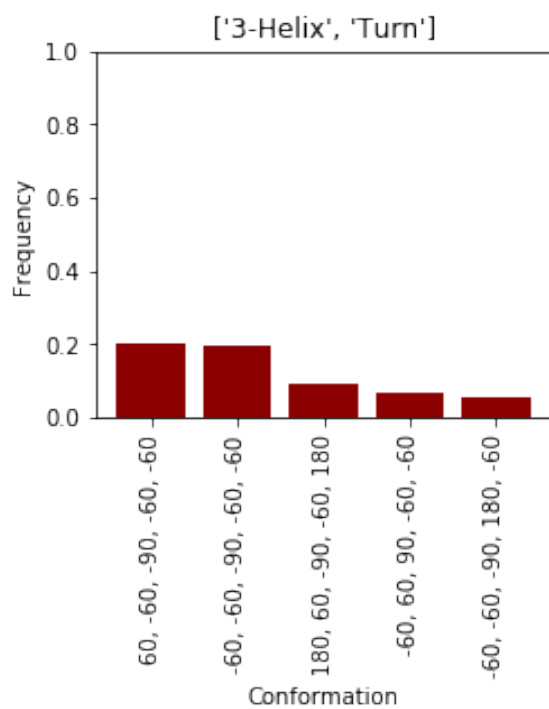
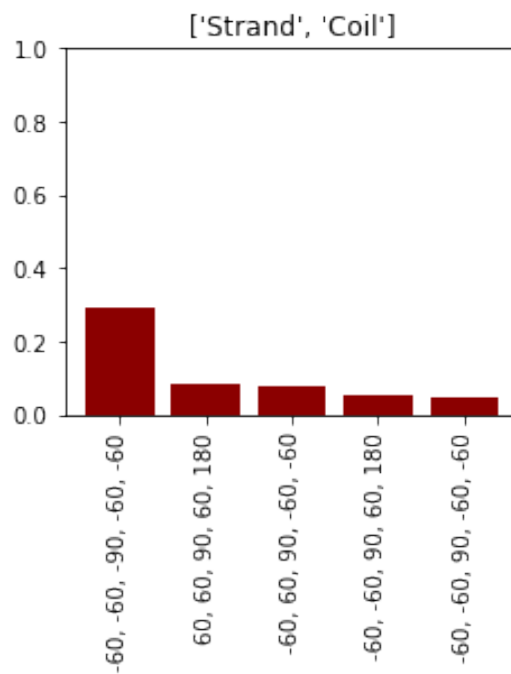
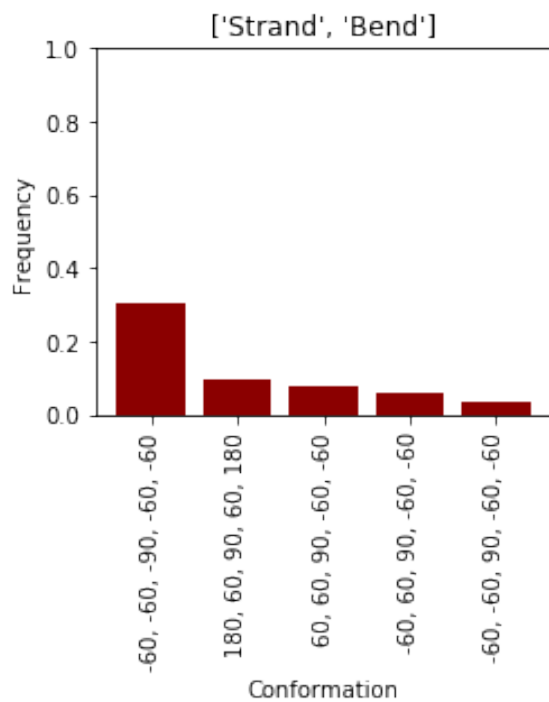


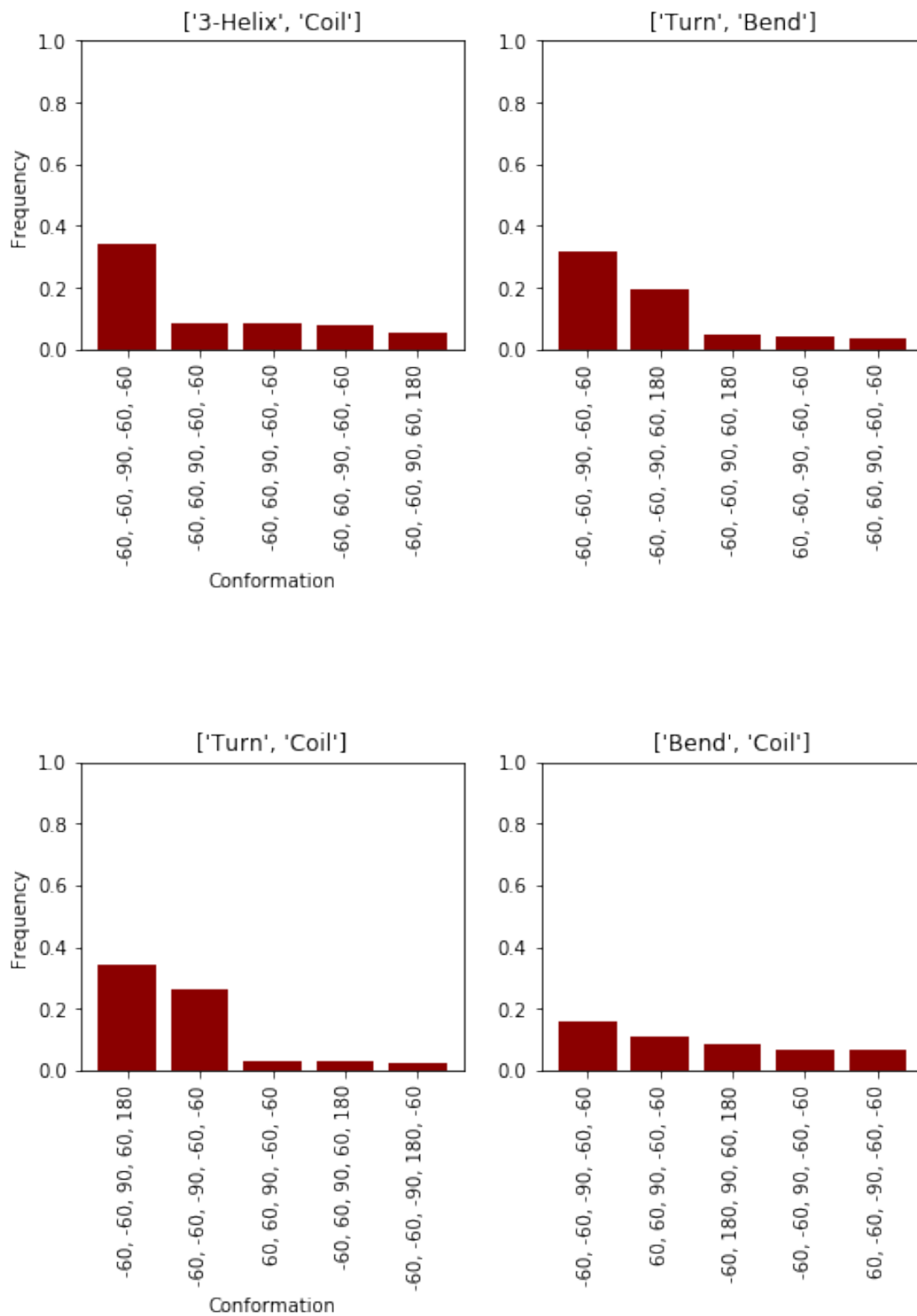


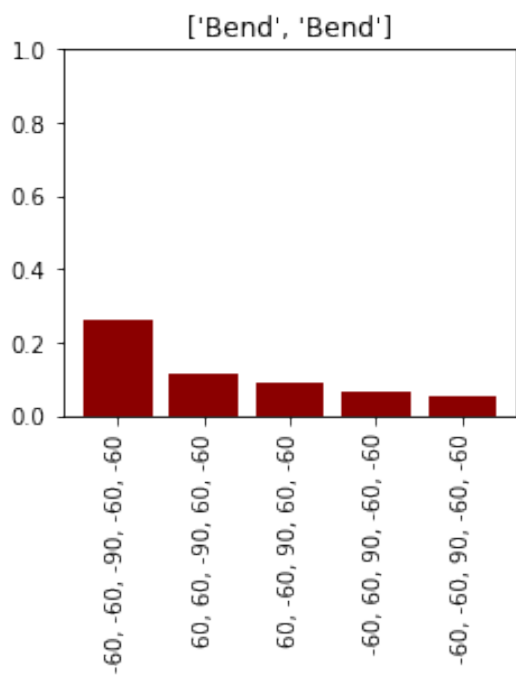
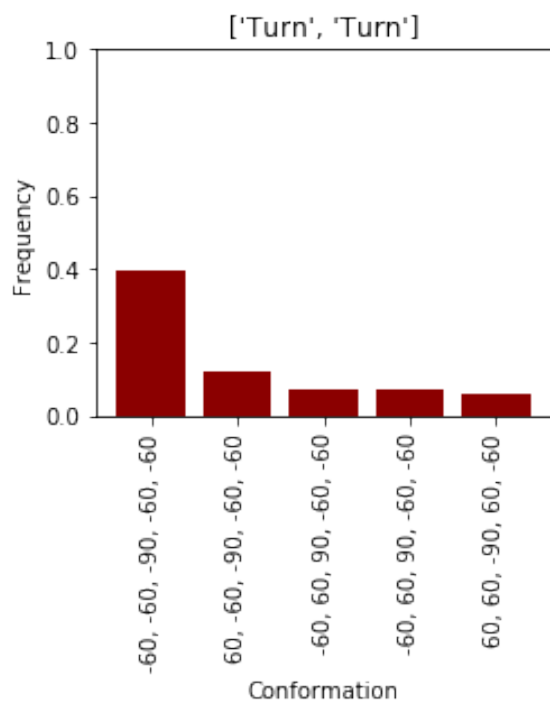
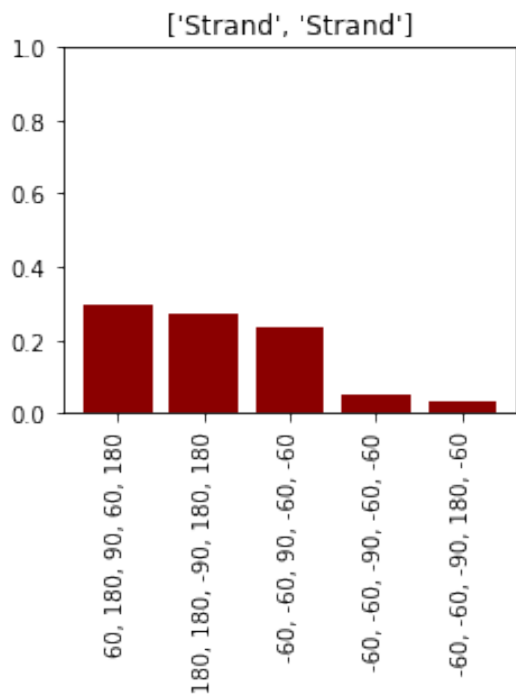
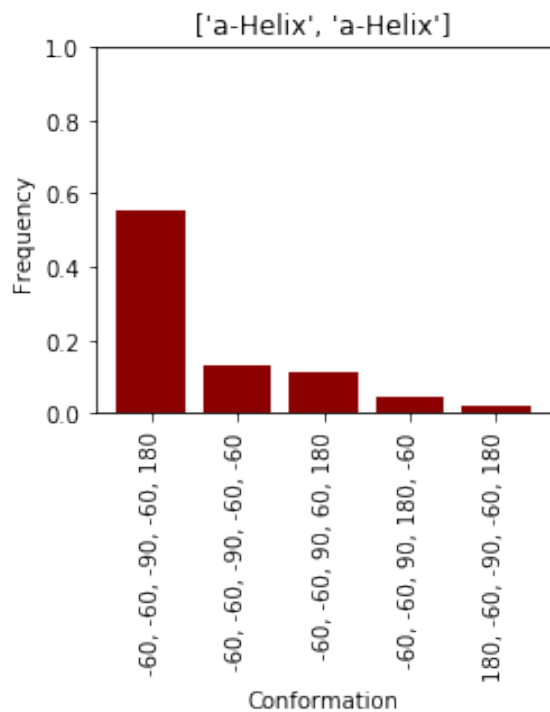




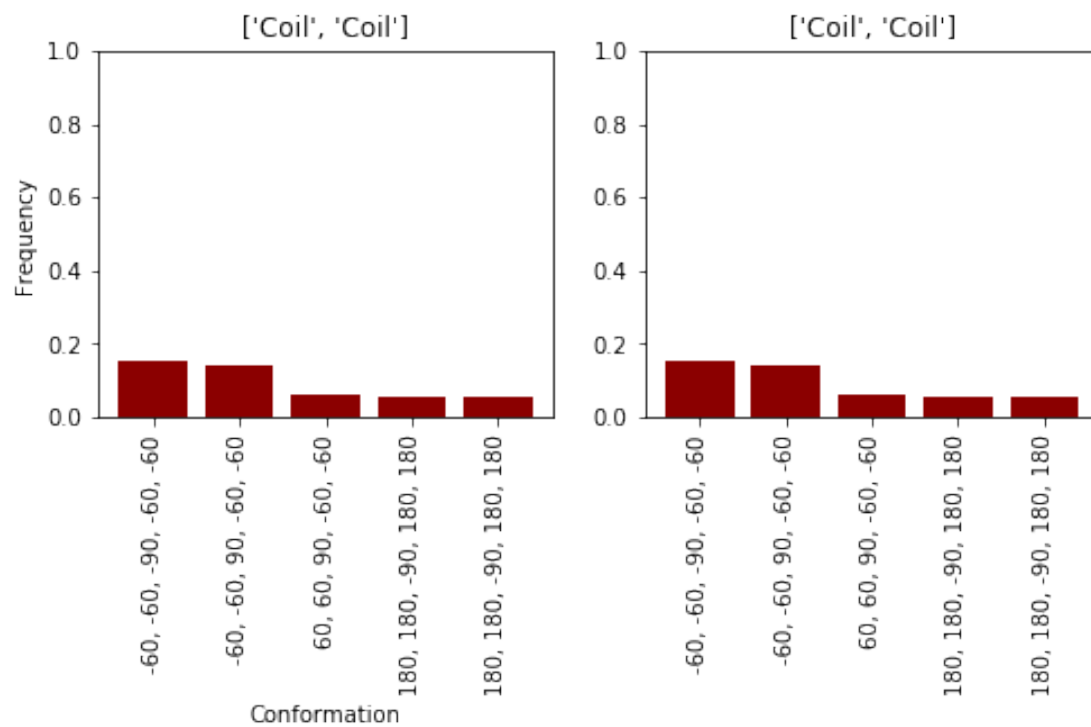


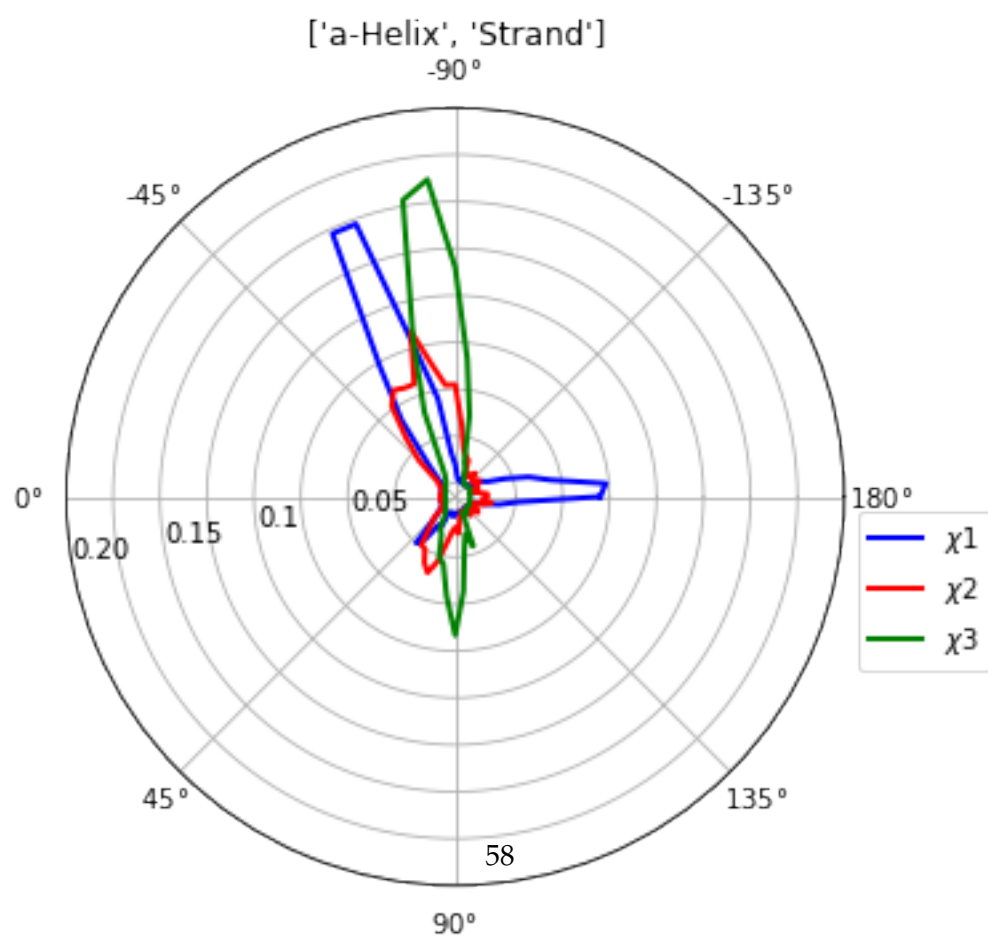
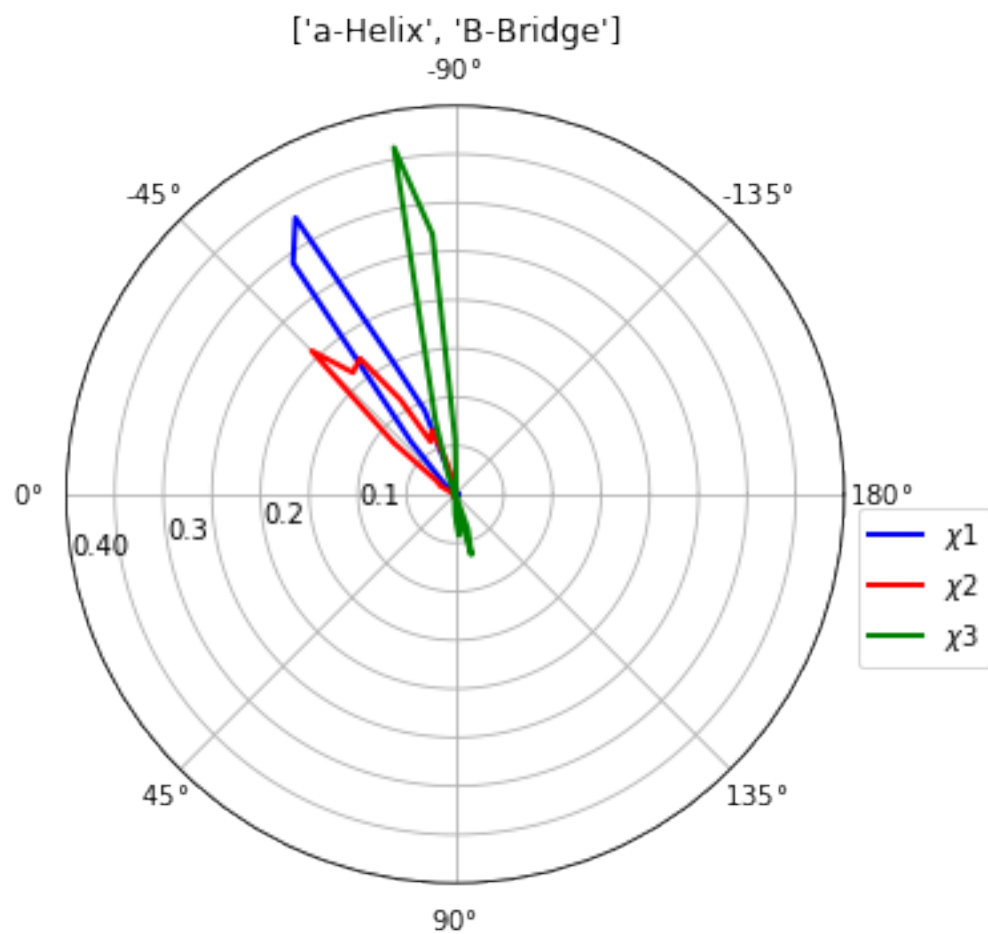








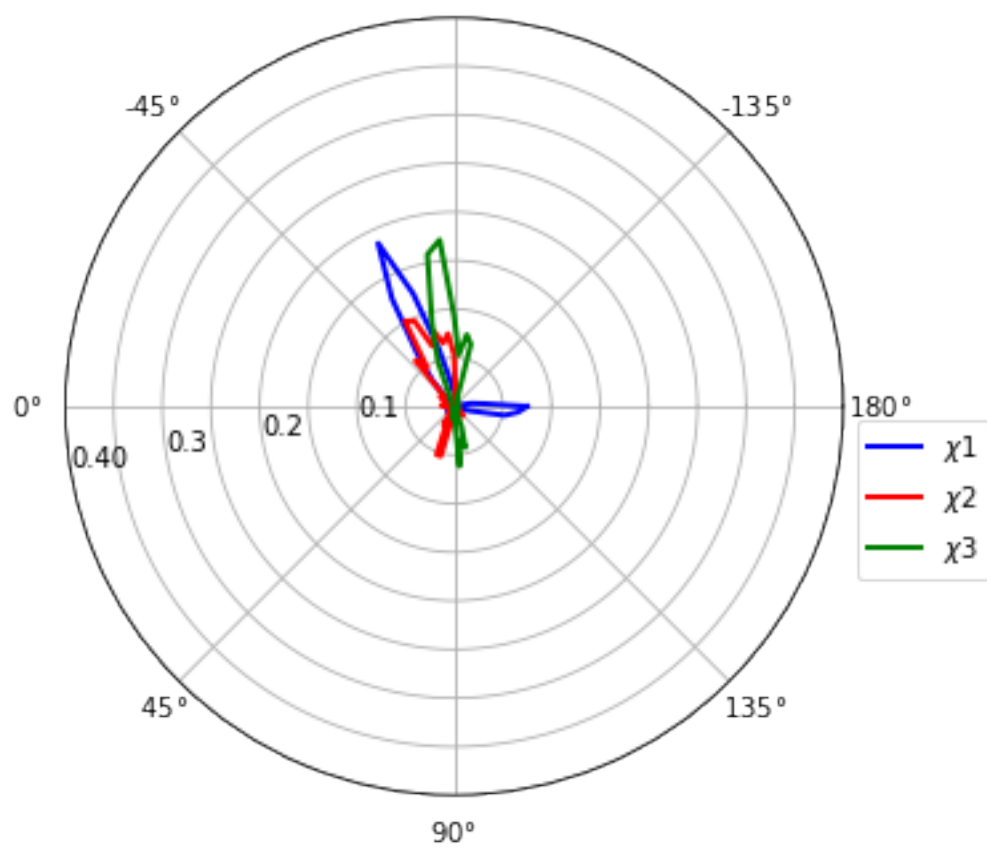






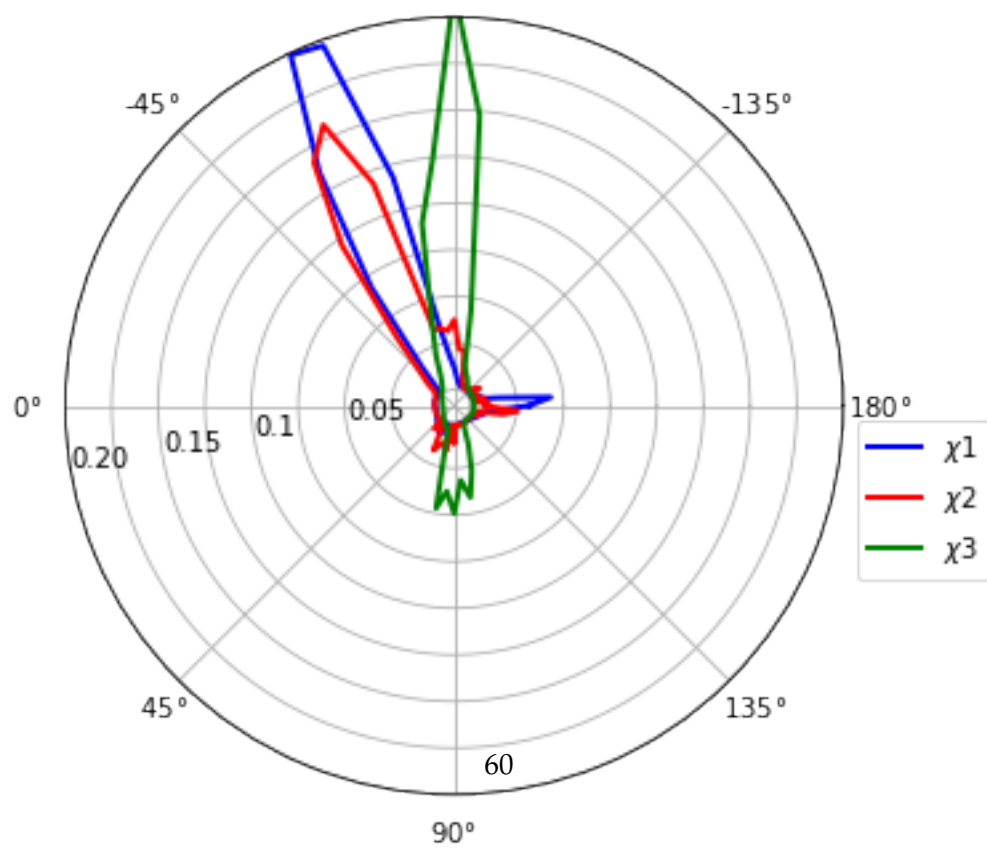
['a-Helix', '3-Helix']

-90°

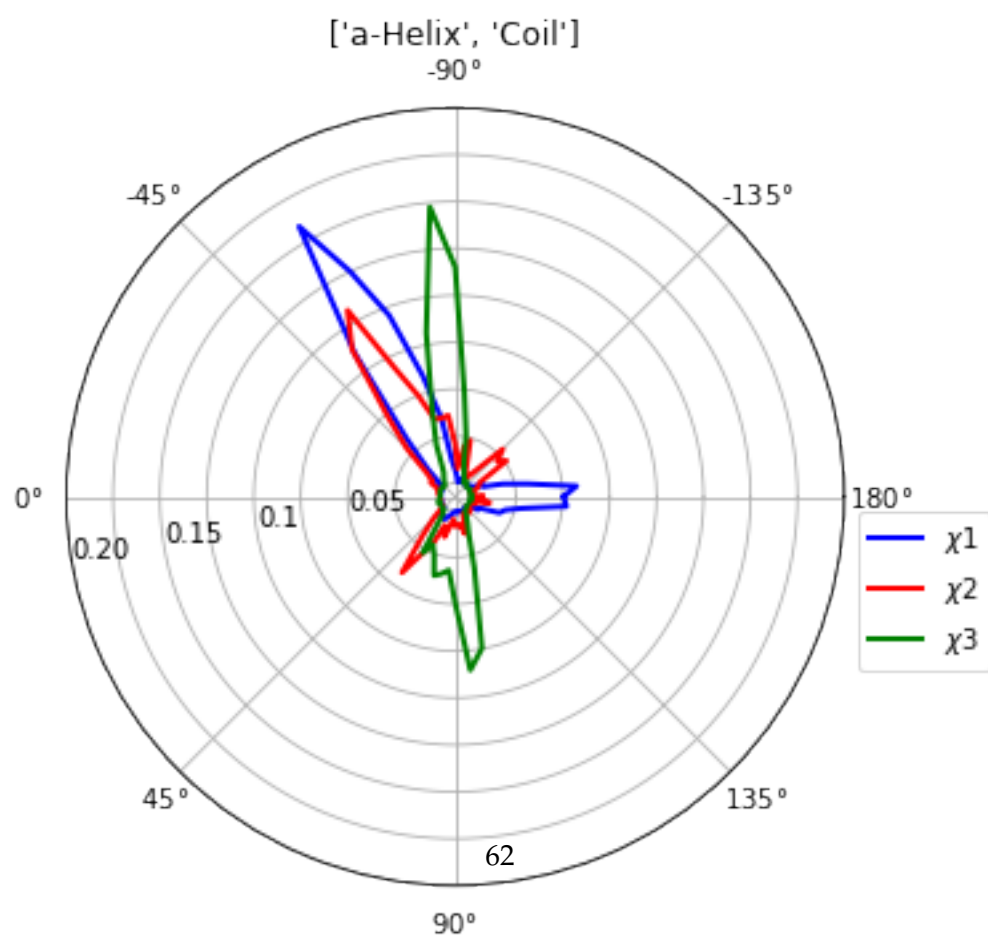
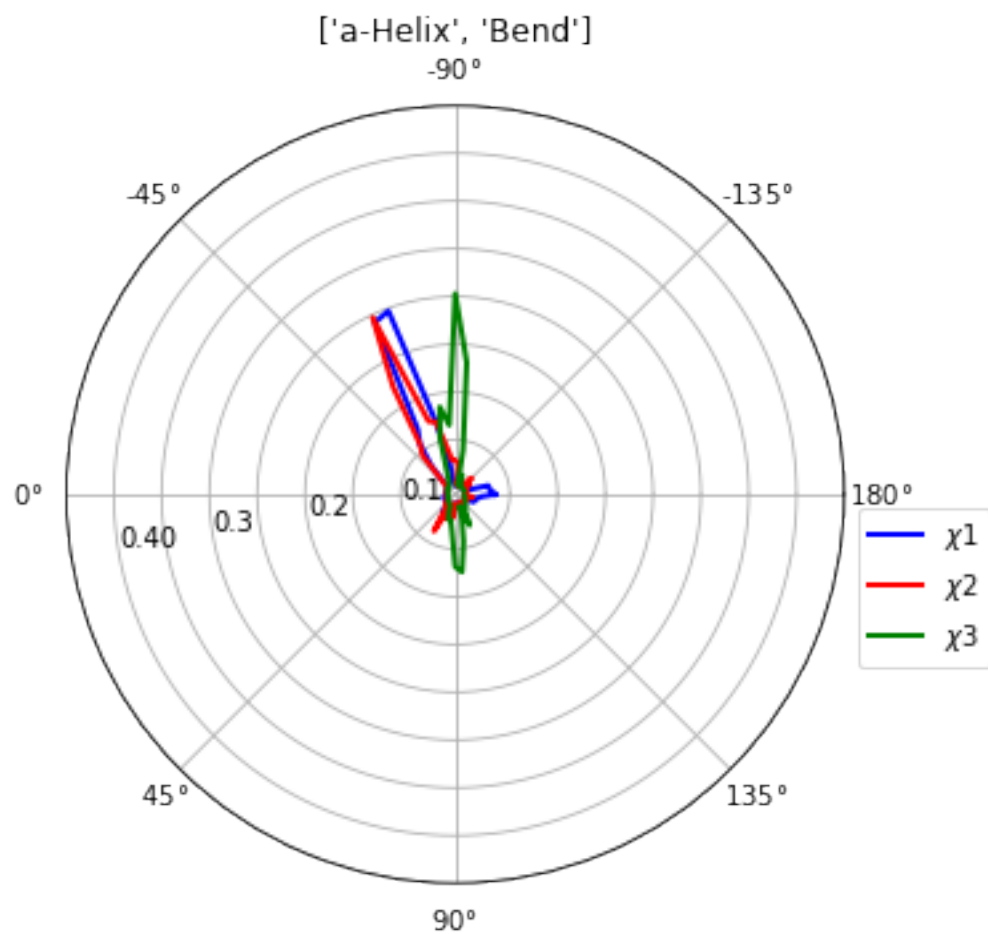


['a-Helix', 'Turn']

-90°

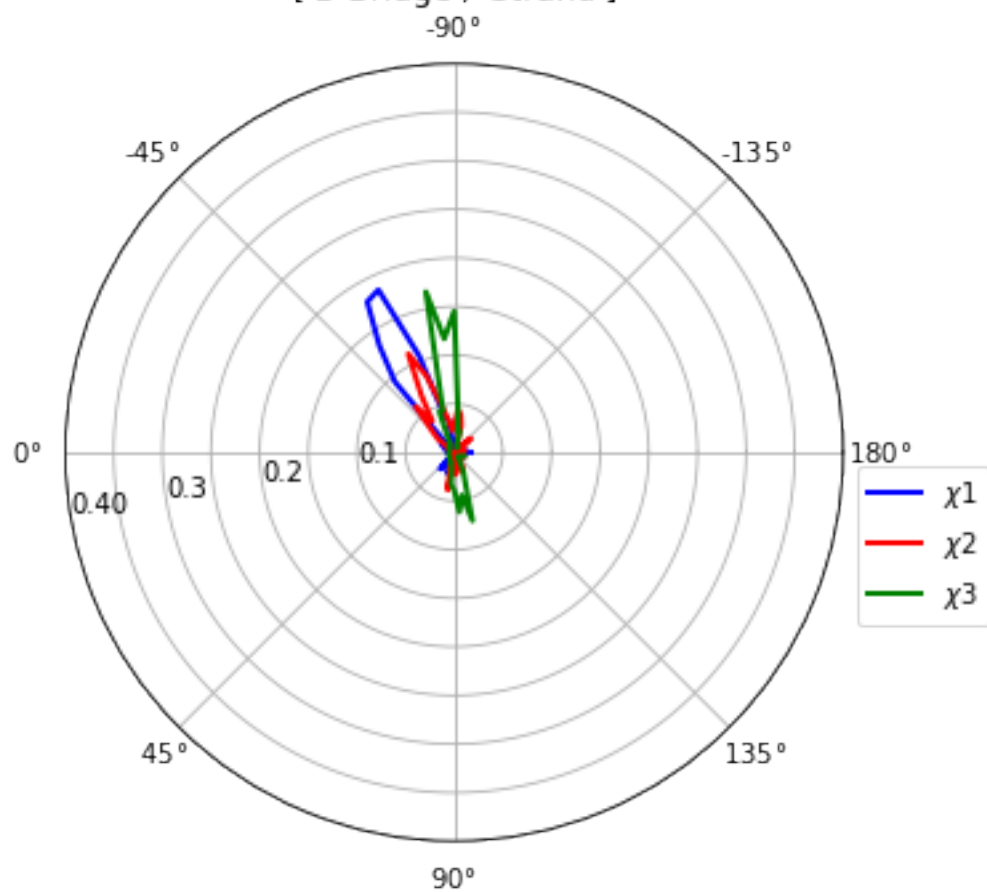




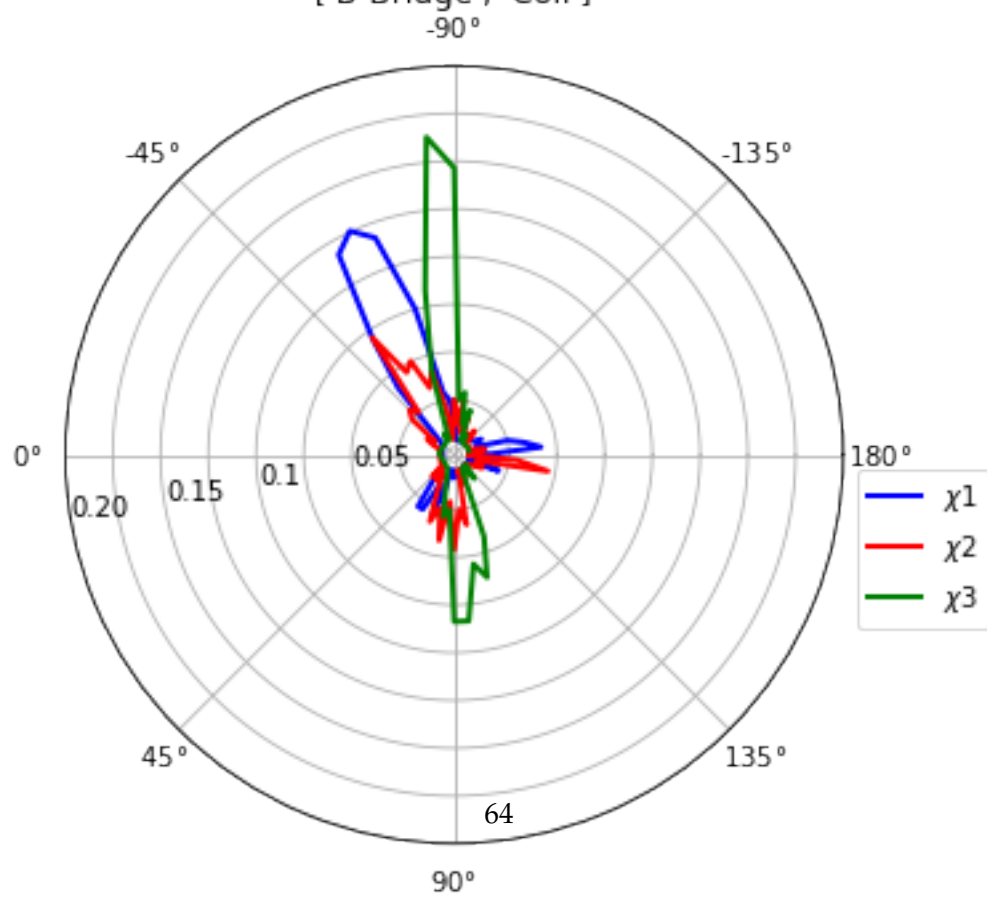




['B-Bridge', 'Strand']

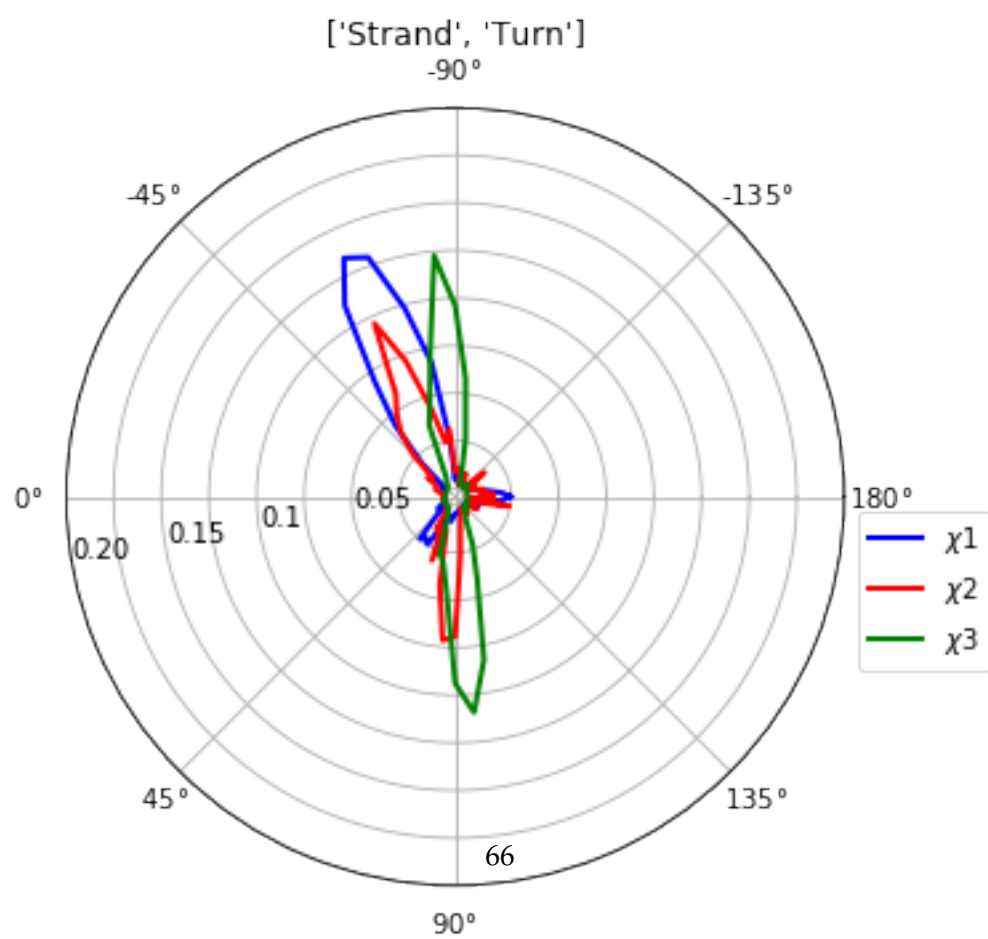
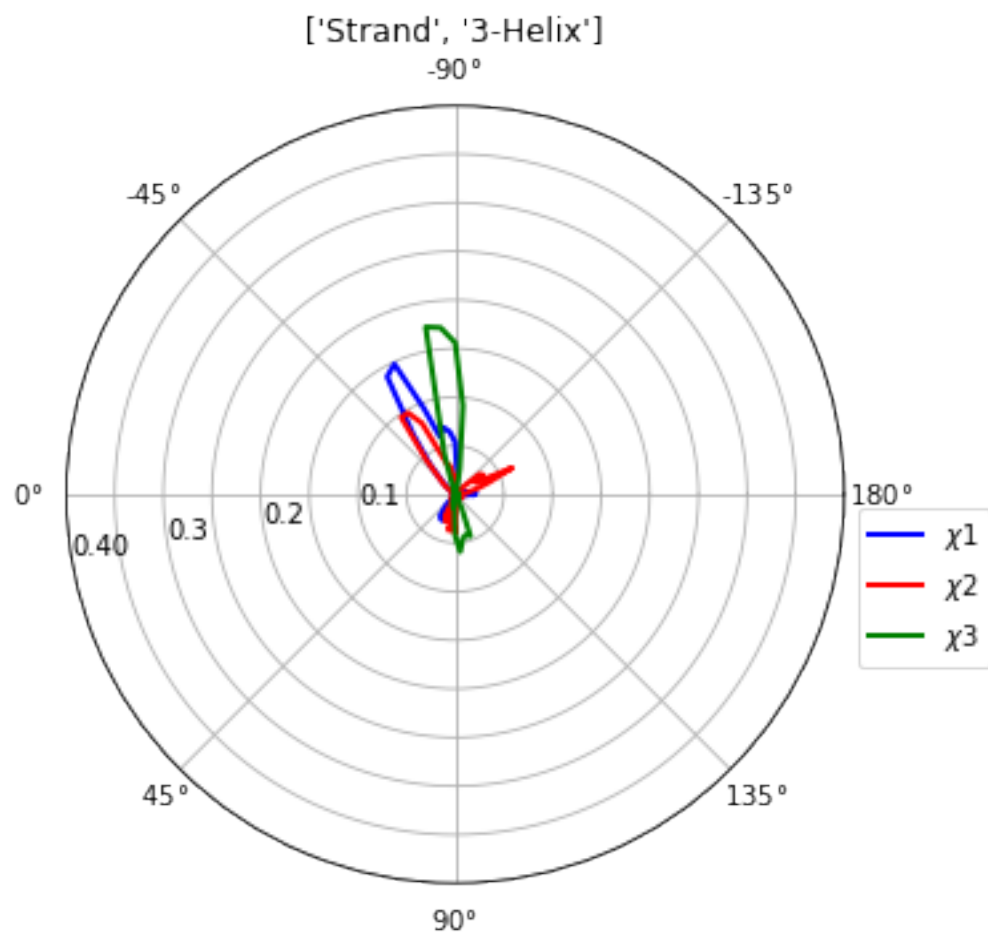


['B-Bridge', 'Coil']

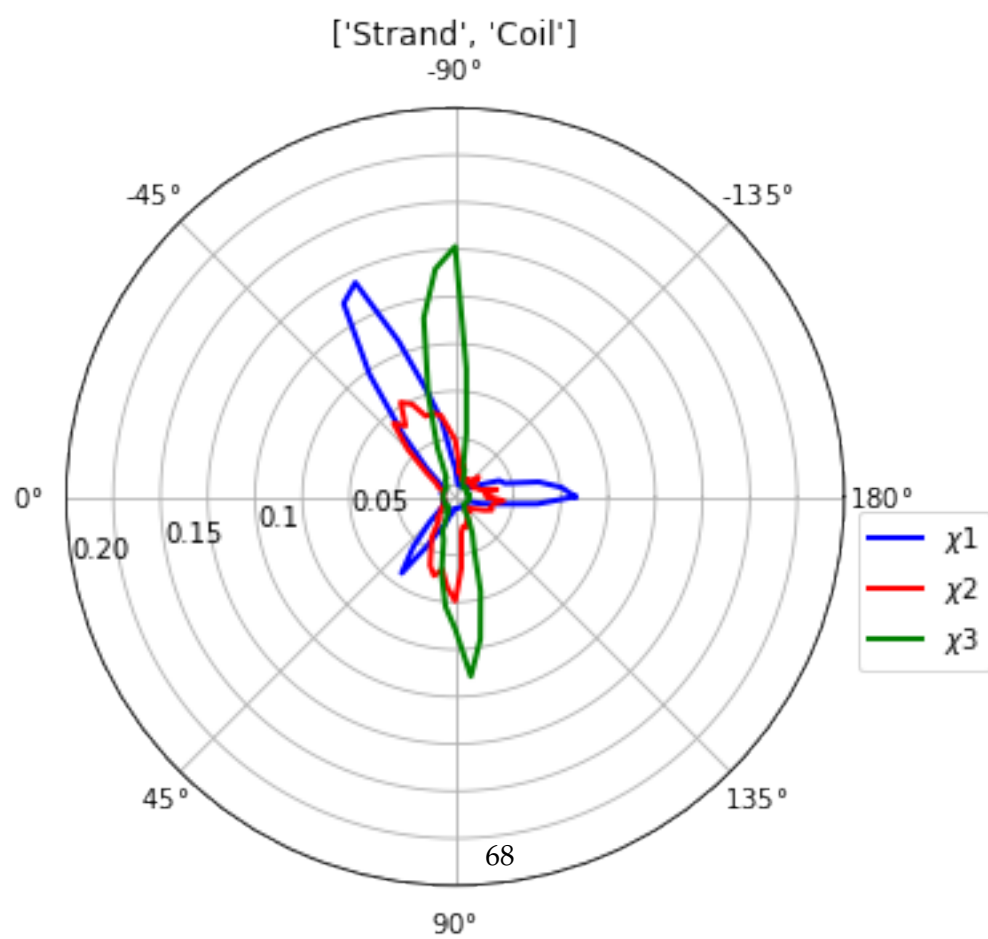
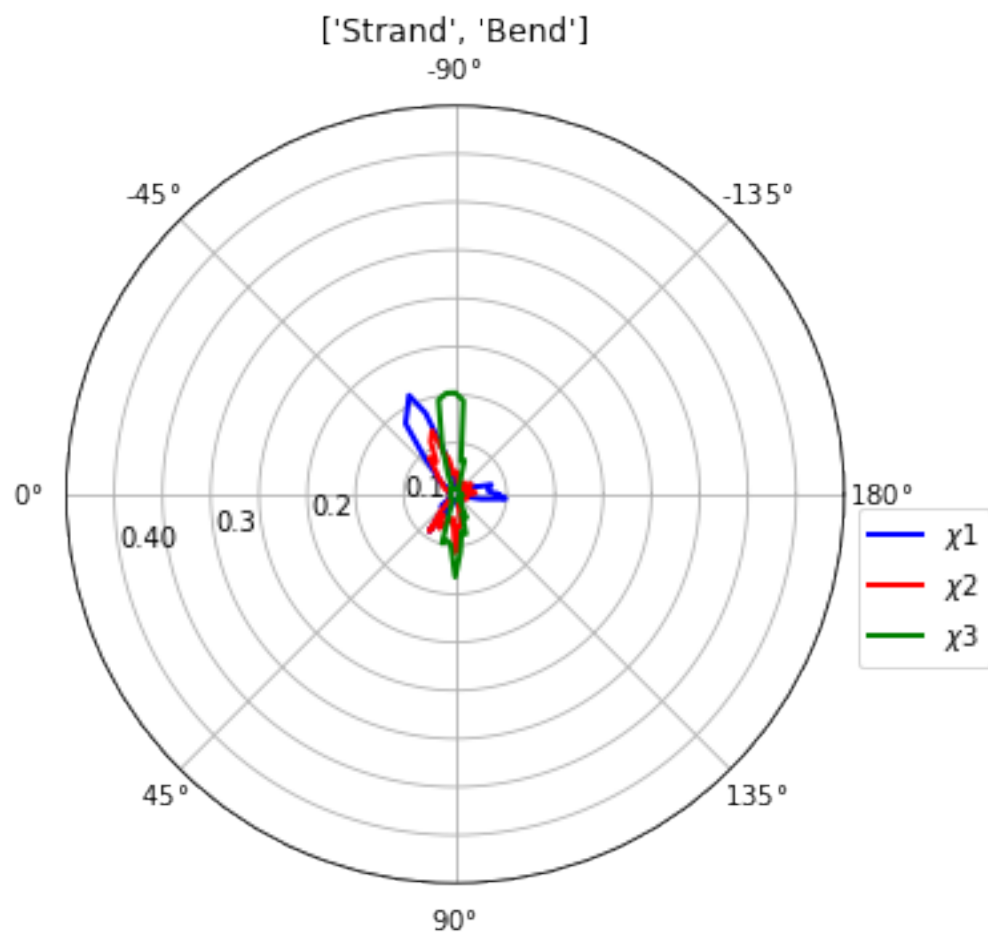








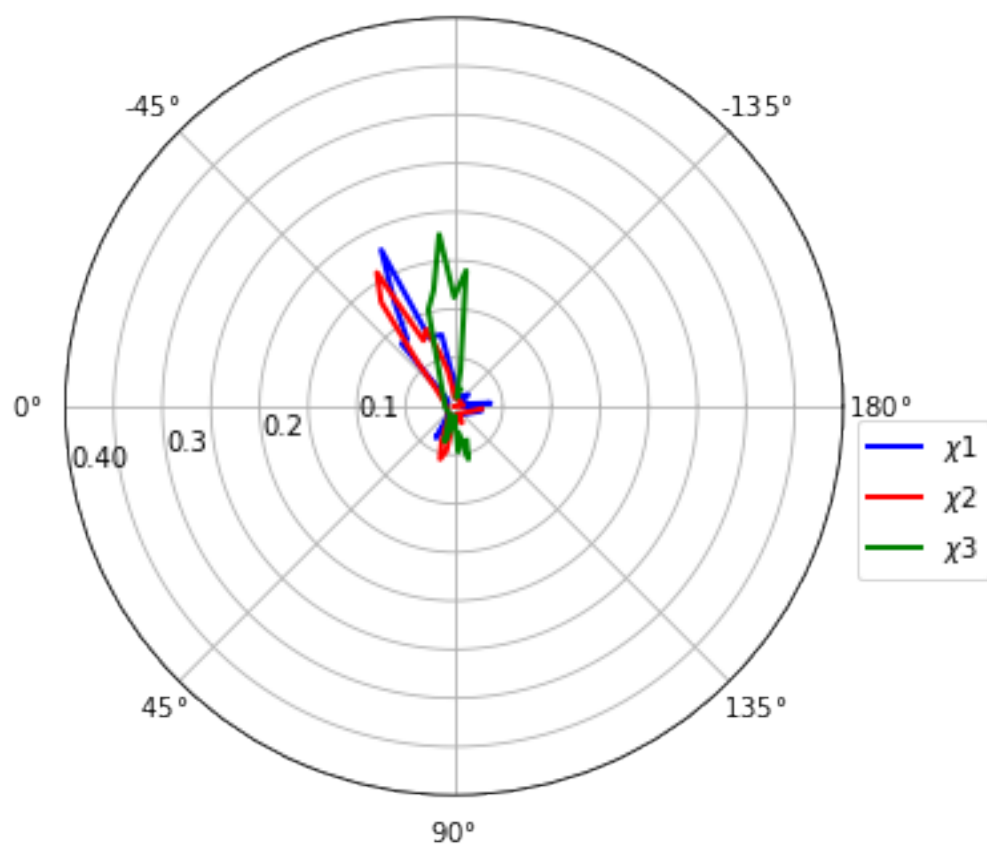






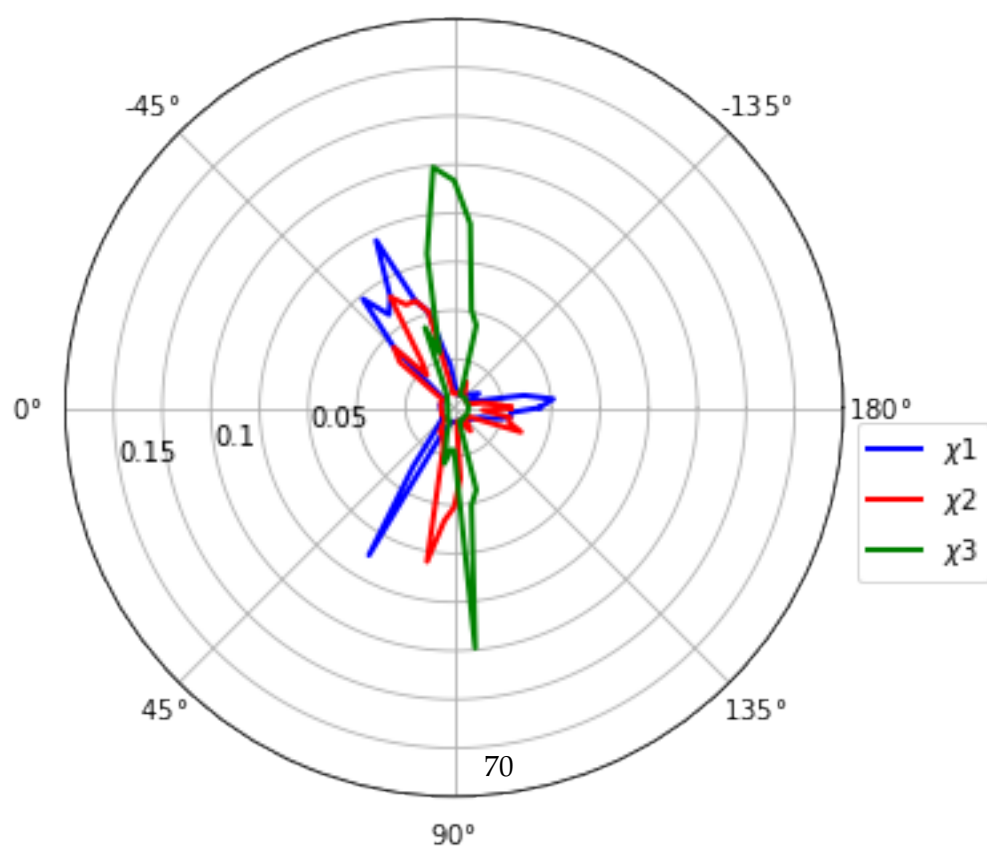
['3-Helix', 'Turn']

-90°



['3-Helix', 'Bend']

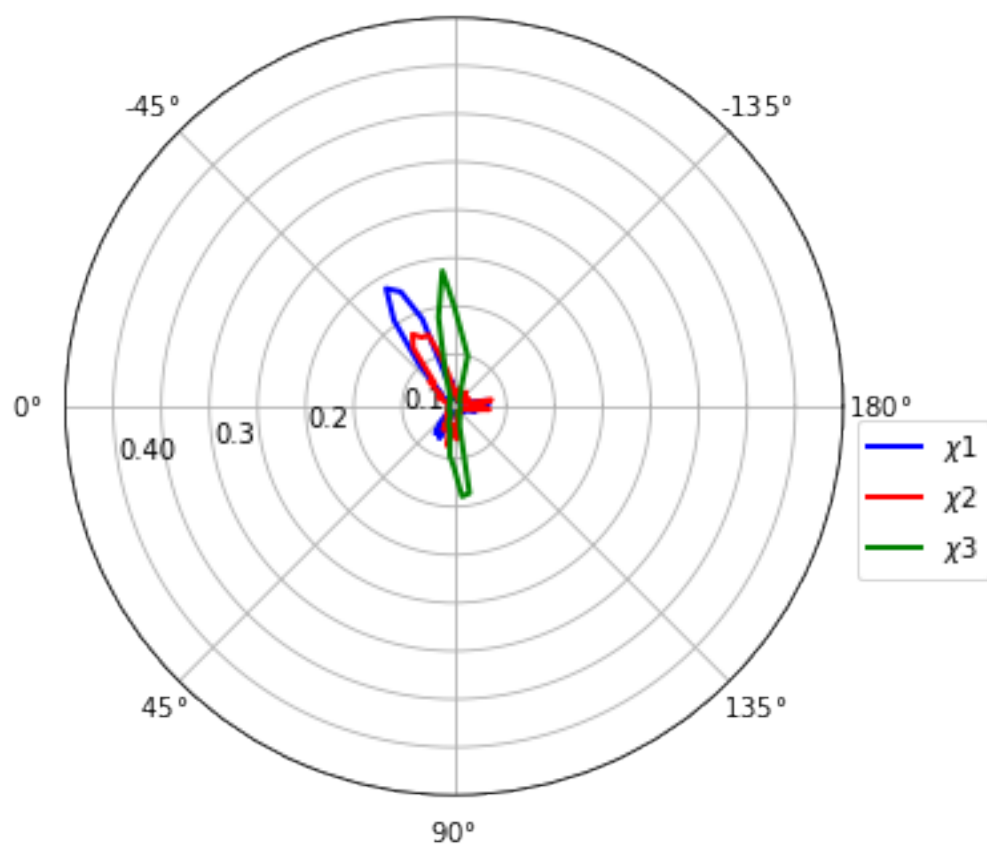
-90°





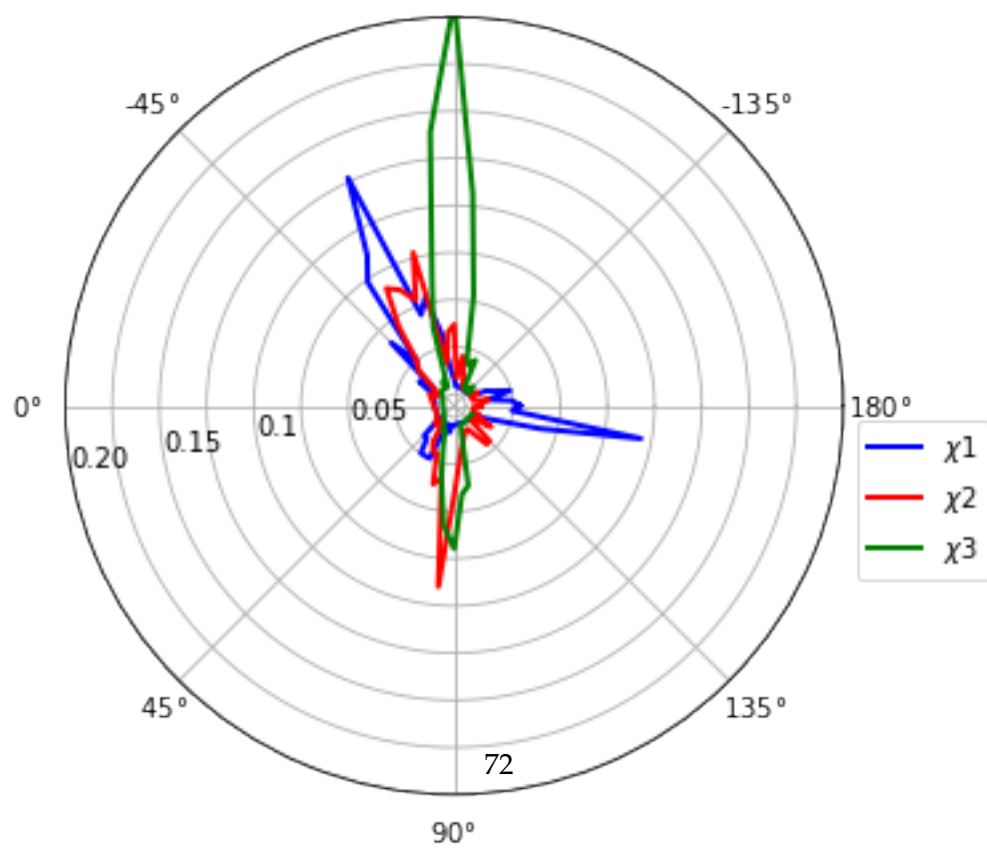
['3-Helix', 'Coil']

-90°



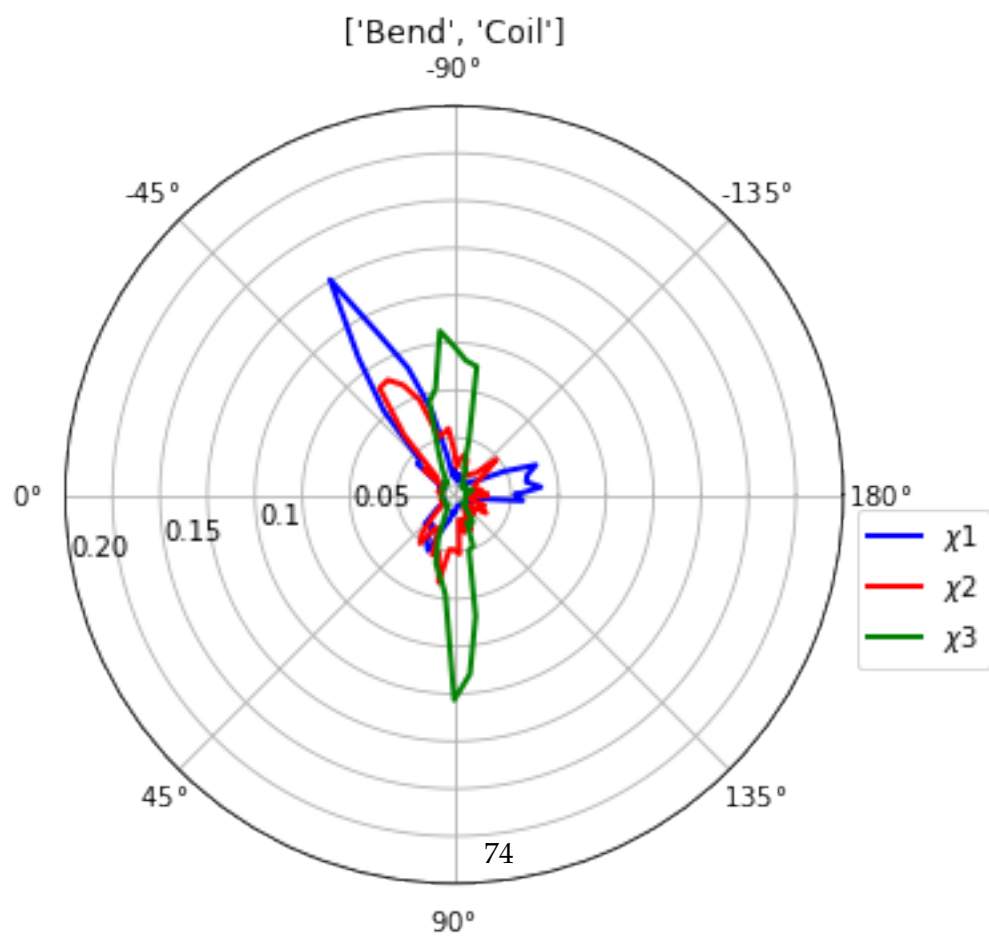
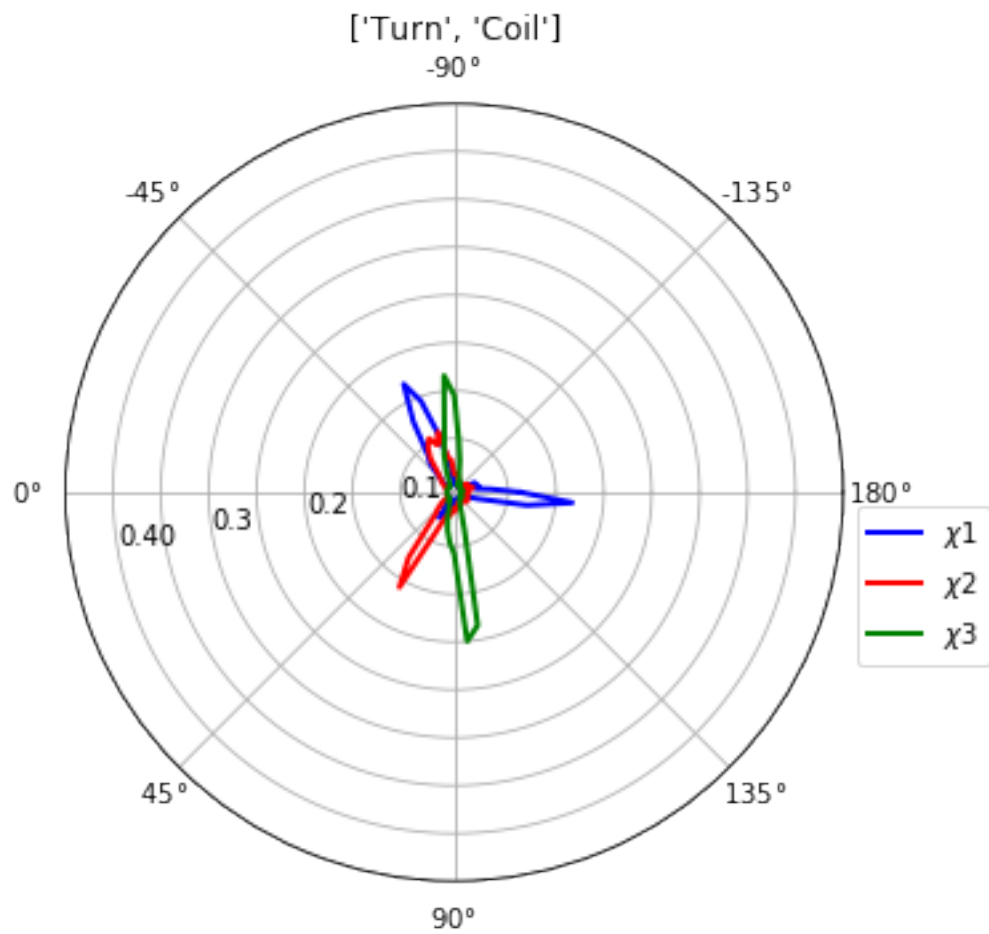
['Turn', 'Bend']

-90°









## **1.8 Dihedral configuraiton by Secondary Structure**

## **1.9 Energy analysis**