Weekly Report(Up until 13th November, 2015)

Anirudh Tiwari

Work Done

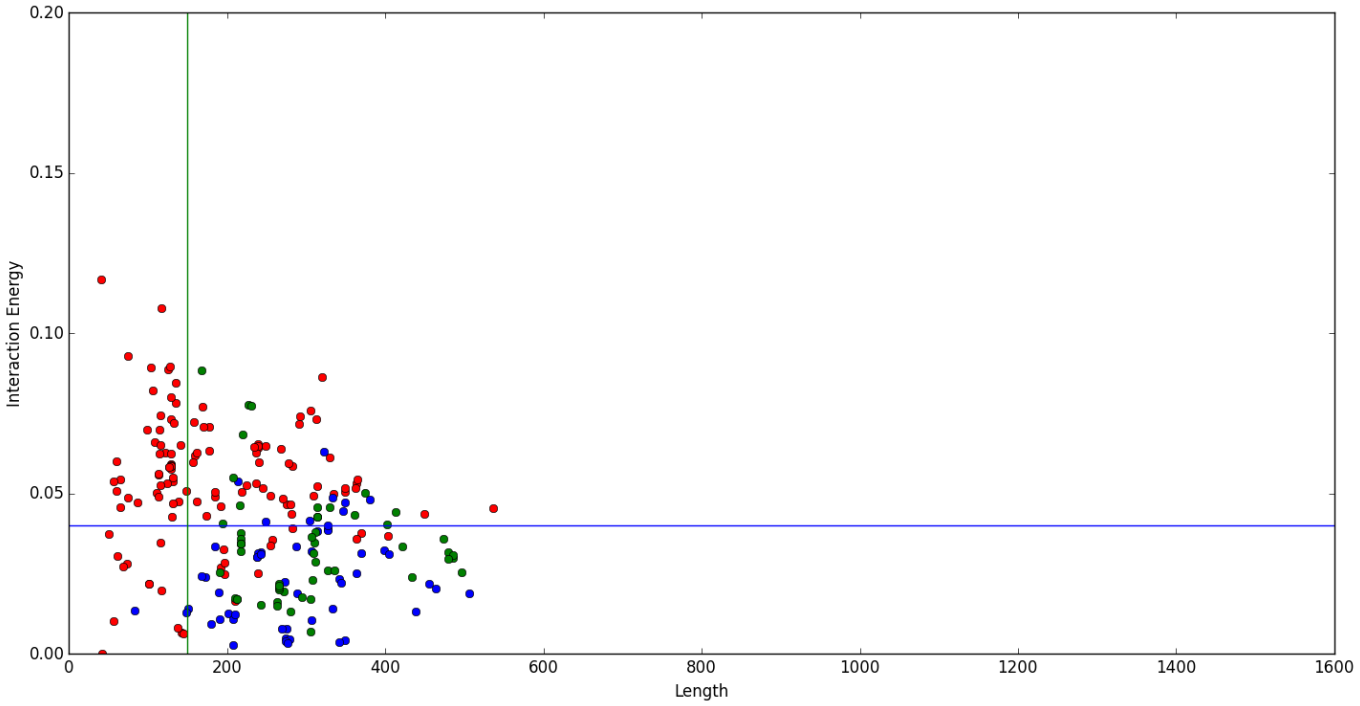
Up until last time, I was trying to classify single domain proteins from two domain proteins. Using the same idea, I tried to classify two domain proteins from three domain, three domain from four domain and so on.

Results

Single Domain vs Two Domain Proteins

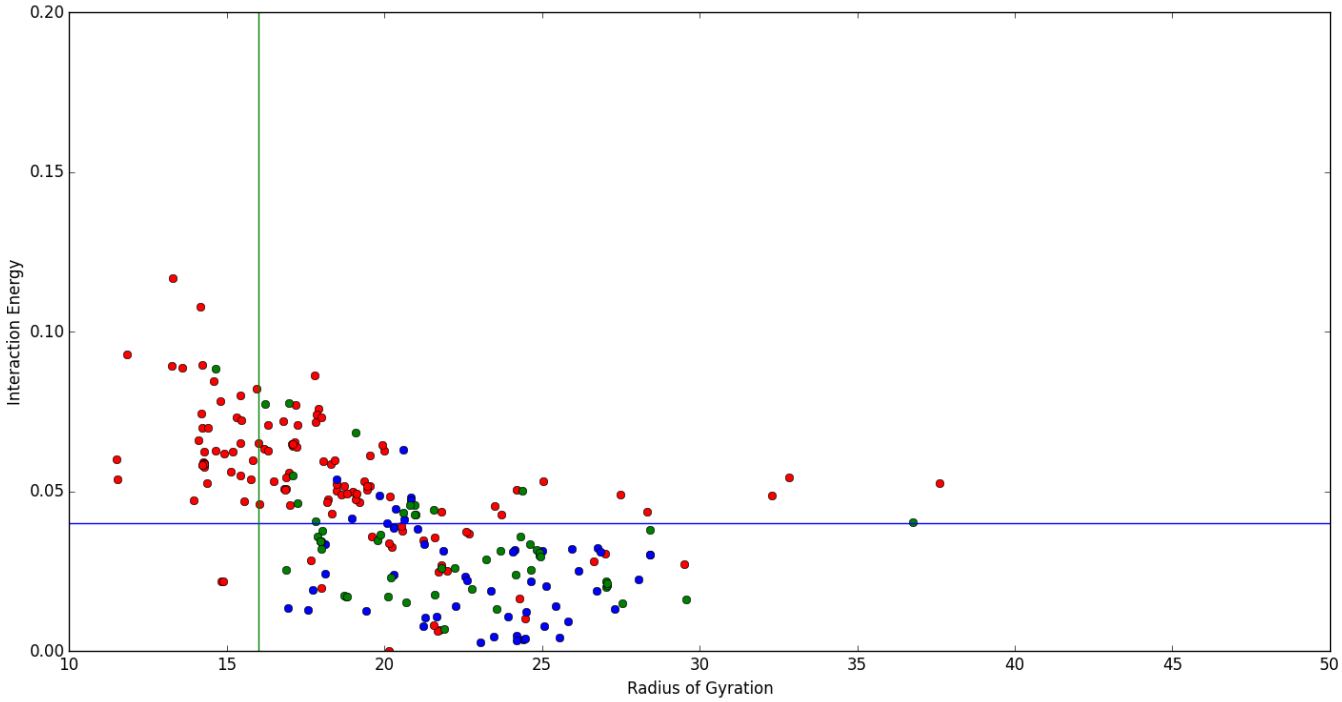
Length vs Interaction Energy

(red dots→ single domain, blue dots→two domain contiguous, green dots→ two domain non-contiguous)



Radius of Gyration vs Interaction Energy

(red dots→ single domain, blue dots→two domain contiguous, green dots→ two domain non-contiguous)

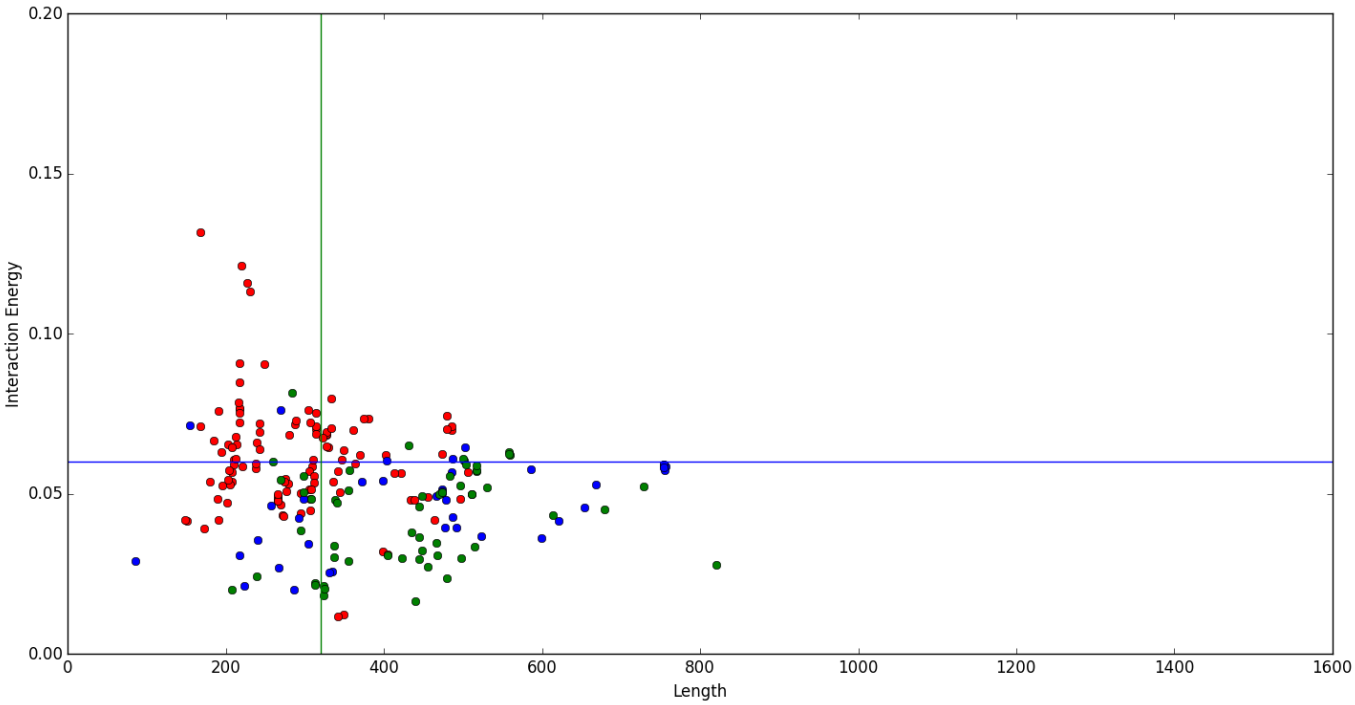


	Length	Interaction Energy	Radius of Gyration	True Positives	False Positives
Single vs Two	>=150	<=0.04	>=16.0	82/108	12/120

Two Domain vs Three Domain Proteins

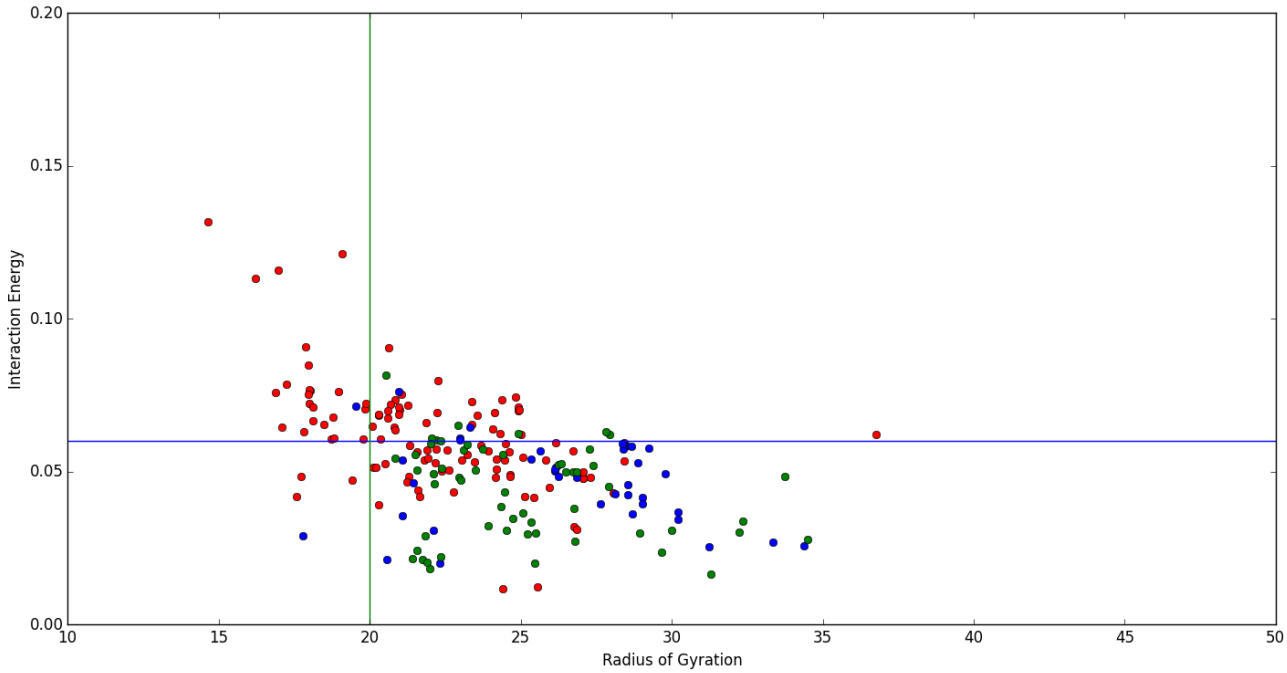
Length vs Interaction Energy

(red dots→ two domain, blue dots→three domain contiguous, green dots→ three domain non-contiguous)



Radius of Gyration vs Interaction Energy

(red dots→ two domain, blue dots→three domain contiguous, green dots→ three domain non-contiguous)

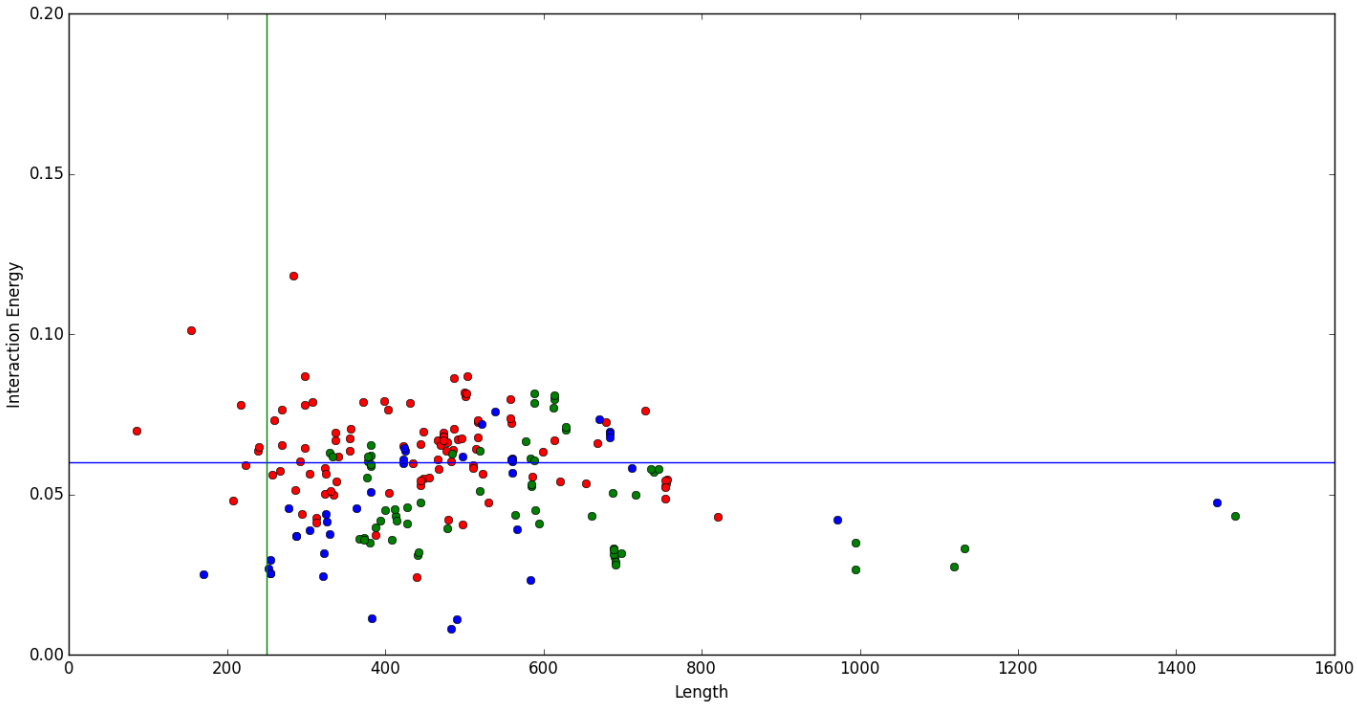


	Length	Interaction Energy	Radius of Gyration	True Positives	False Positives
Two vs Three	>=320	<=0.06	>=20.0	65/97	16/112

Three Domain vs Four Domain Proteins

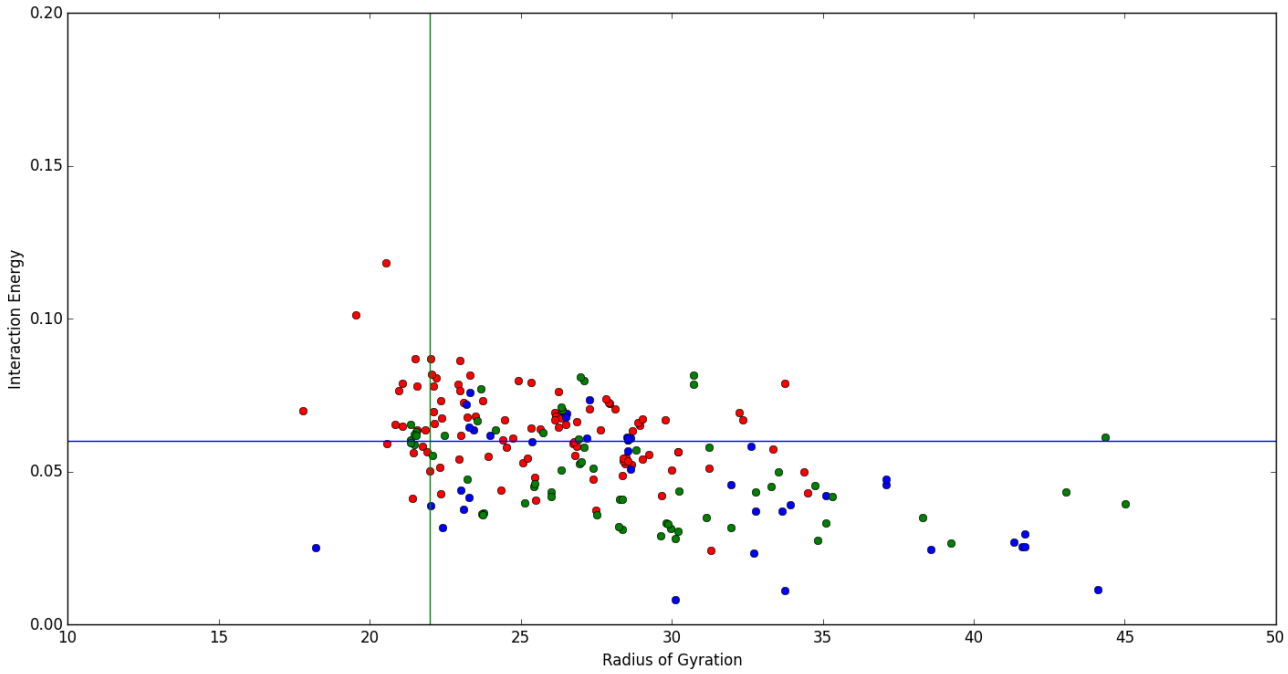
Length vs Interaction Energy

(red dots→ three domain, blue dots→four domain contiguous, green dots→ four domain non-contiguous)



Radius of Gyration vs Interaction Energy

(red dots→ three domain, blue dots→four domain contiguous, green dots→ four domain non-contiguous)

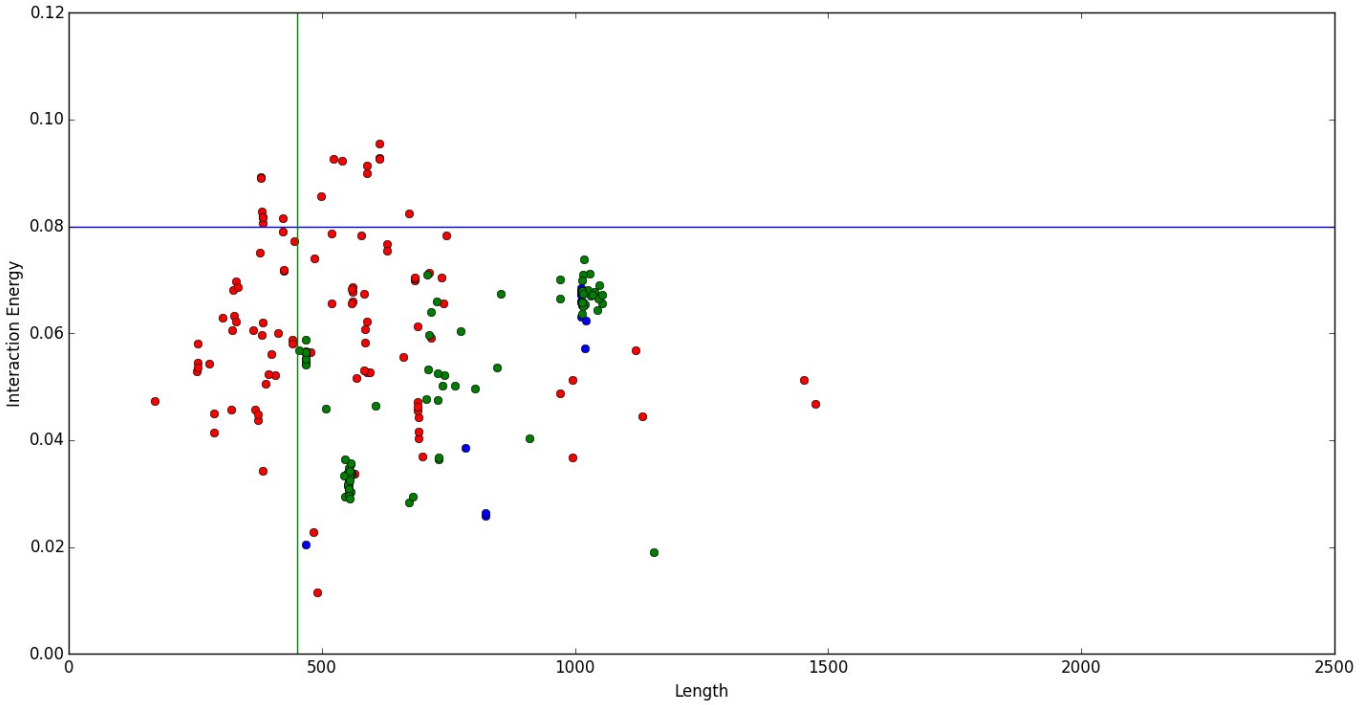


	Length	Interaction Energy	Radius of Gyration	True Positives	False Positives
Three vs Four	>=250	<=0.06	>=22.0	67/102	33/98

Four Domain vs Five Domain Proteins

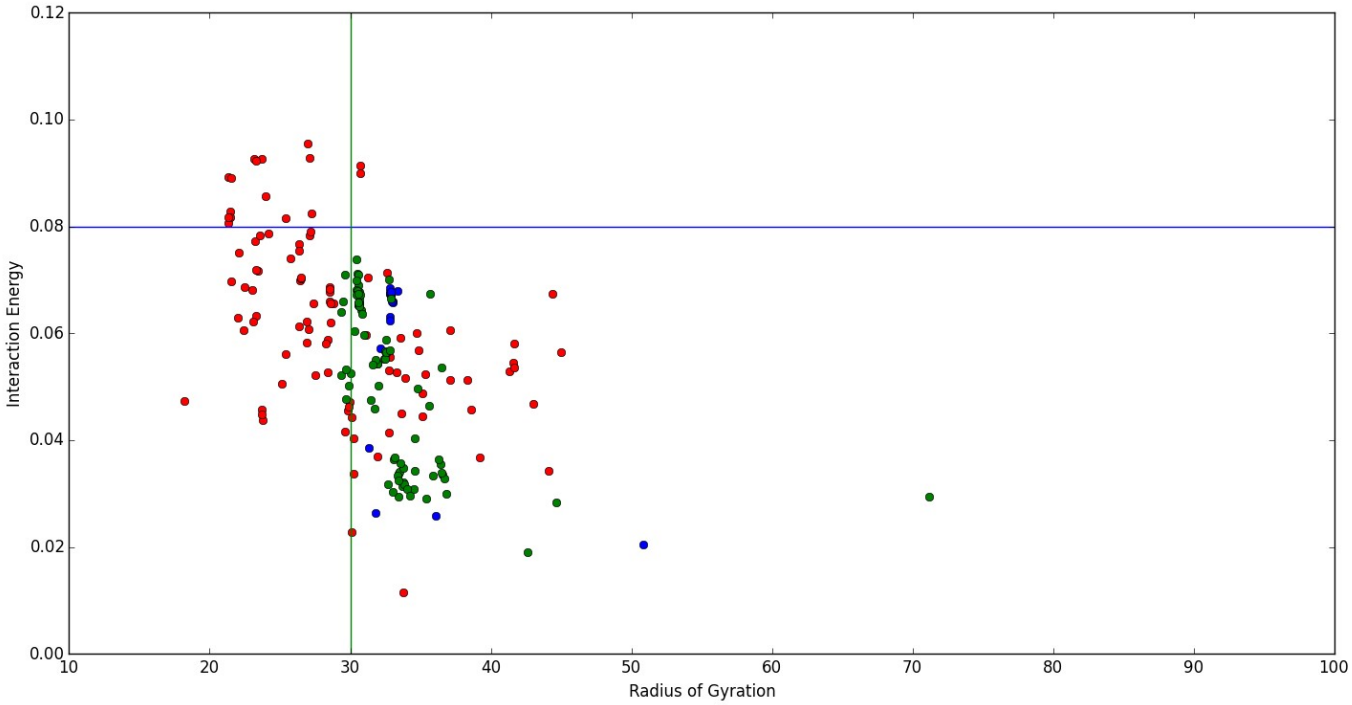
Length vs Interaction Energy

(red dots→ four domain, blue dots→five domain contiguous, green dots→ five domain non-contiguous)



Radius Of Gyration vs Interaction Energy

(red dots→ four domain, blue dots→five domain contiguous, green dots→ five domain non-contiguous)



	Length	Interaction Energy	Radius of Gyration	True Positives	False Positives
Four vs Five	>=450	<=0.08	>=30.0	93/100	22/98

Summary Of Results

	Length	Interaction Energy	Radius of Gyration	True Positives	False Positives
Single vs Two	>=150	<=0.04	>=16.0	82/108	12/120
Two vs Three	>=320	<=0.06	>=20.0	65/97	16/112
Three vs Four	>=250	<=0.06	>=22.0	67/102	33/98
Four vs Five	>=450	<=0.08	>=30.0	93/100	22/98

Next Steps

- These results are based on CATH classification only, I'll try to incorporate SCOP information also to check whether it improves my result or not.
- This dataset of proteins is generated by me, I'll run it on Holland et al bench marked data set.