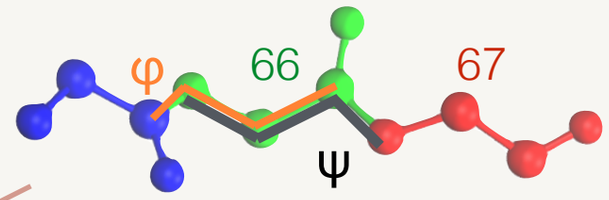
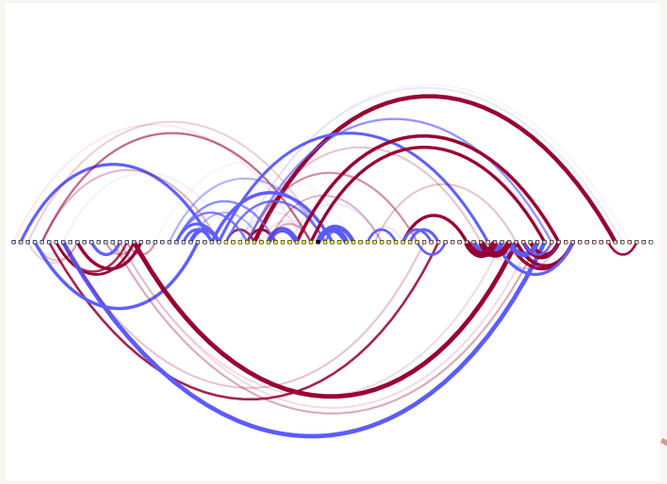
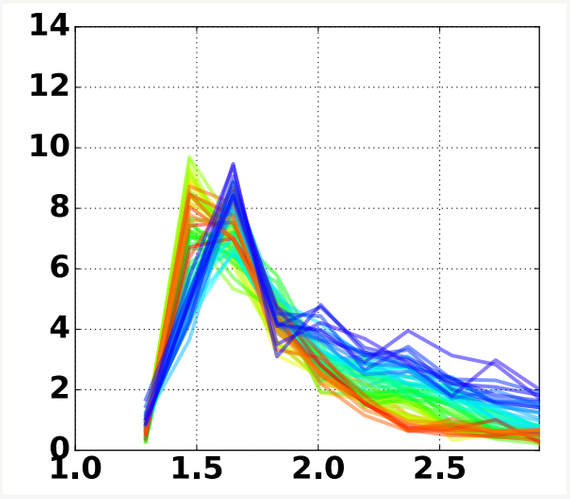
Mechanism underlying conformational effects of a disease-associated hydrophobic-to-hydrophobic substitution on an intrinsically disordered region







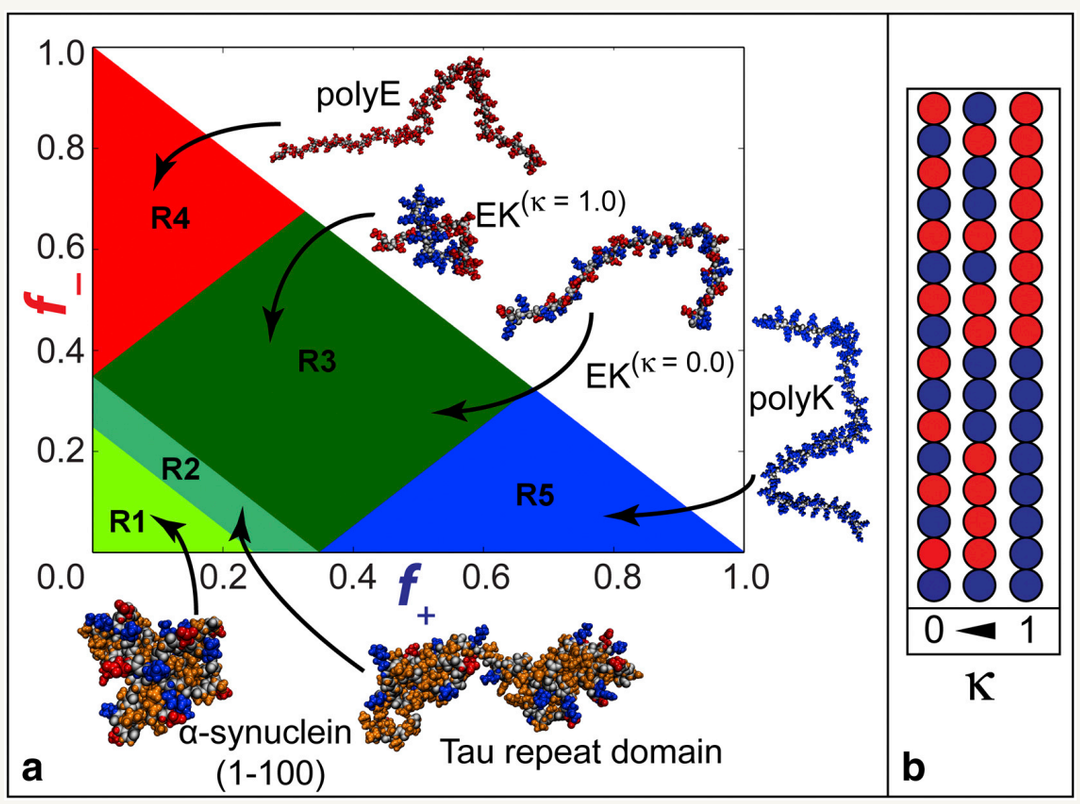
Presented by:
Ruchi Lohia

Physiological significance of intrinsically disordered proteins (IDPs)

- > 33 % of eukaryotic proteins have long (>30 residues) disordered regions
- Involved in critical biological functions including transcriptional activation and intracellular signaling
- Can undergo coupled folding and binding
- Can bind multiple partners with high specificity and low affinity
- > 20 % of missense disease mutations are located in disordered regions
- Implicated in various diseases, including neuro-degeneration and cancer

Sequence ensemble relationship for IDP's based on polymer physics

- >75% the IDP's are polyampholytes and shows the sequence ensemble relationship as predicted by Pappu's lab



(Das et al., PNAS, 2013)

f_+ fraction of positively charged residues

f_- fraction of negatively charged residues

κ sequence distribution of oppositely charged residues

negatively charged

positively charged

- Weak polyampholytes & polyelectrolytes: Globules & tadpoles
- Janus sequences: Collapsed or expanded - context dependent
- Strong polyampholytes: Coils, hairpins, & chimeras
- Negatively charged strong polyelectrolytes: Swollen coils
- Positively charged strong polyelectrolytes: Swollen coils

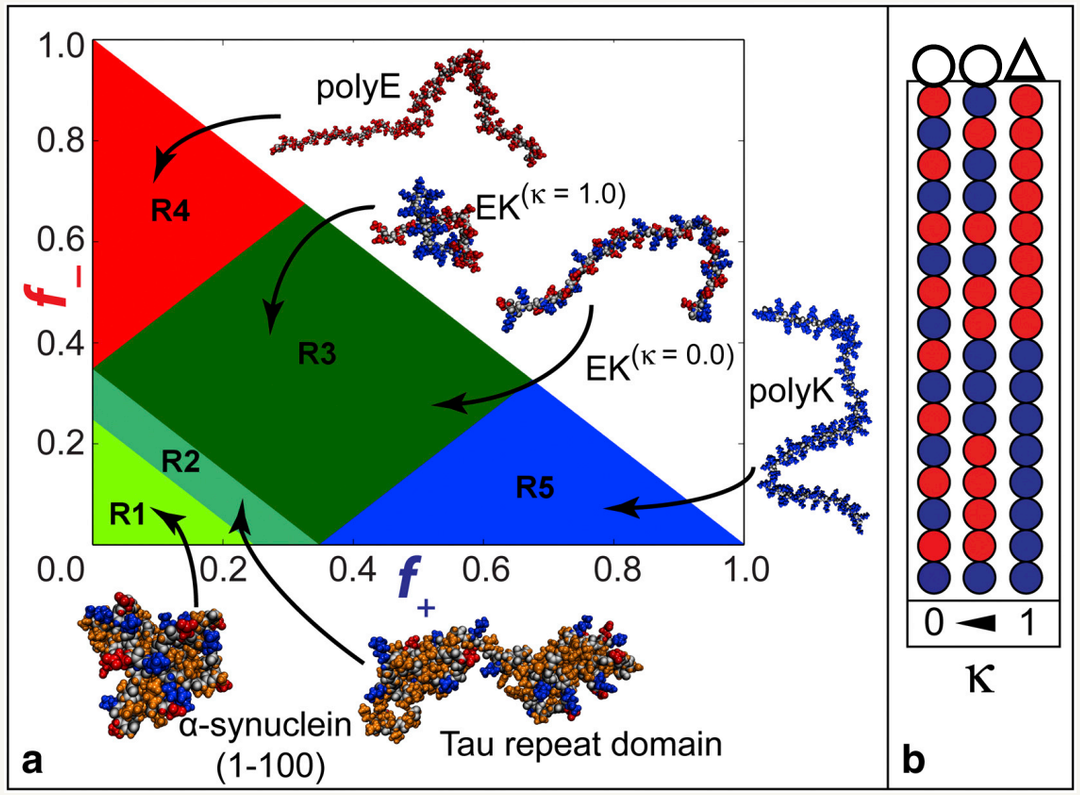
Challenges in predicting IDP's sequence ensemble relationship

- ~70% of the IDP's fall in the R1,R2 region
- The current theory deals with sequence composition and sequence patterning and is completely based on charged residues
- It does not deals with specific residues; for e.g. different polar side-chains will have different effects on the conformational properties and solubility profiles of IDPs

f_+ fraction of positively charged residues

f_- fraction of negatively charged residues

κ sequence distribution of oppositely charged residues



(Das et al., PNAS, 2013)

negatively charged

positively charged

uncharged residues

- Weak polyampholytes & polyelectrolytes: Globules & tadpoles
- Janus sequences: Collapsed or expanded - context dependent
- Strong polyampholytes: Coils, hairpins, & chimeras
- Negatively charged strong polyelectrolytes: Swollen coils
- Positively charged strong polyelectrolytes: Swollen coils

Effect of mutations on IDP's conformational ensemble

- > 20 % of missense disease mutations are located in disordered regions

(Vacic et al.,PLoS Comput Bio, 2013)

- >10% of these mutations are hydrophobic to hydrophobic mutations

Secondary Structure

- Charged residue mutations can increase or decrease residual helicity
- Charged residue mutations can modulate the local helicity by either forming or breaking salt-bridge or by introducing a helix breaking residues (glycine, proline)

(Ganguly et al.,PLoS Comput Bio, 2015)

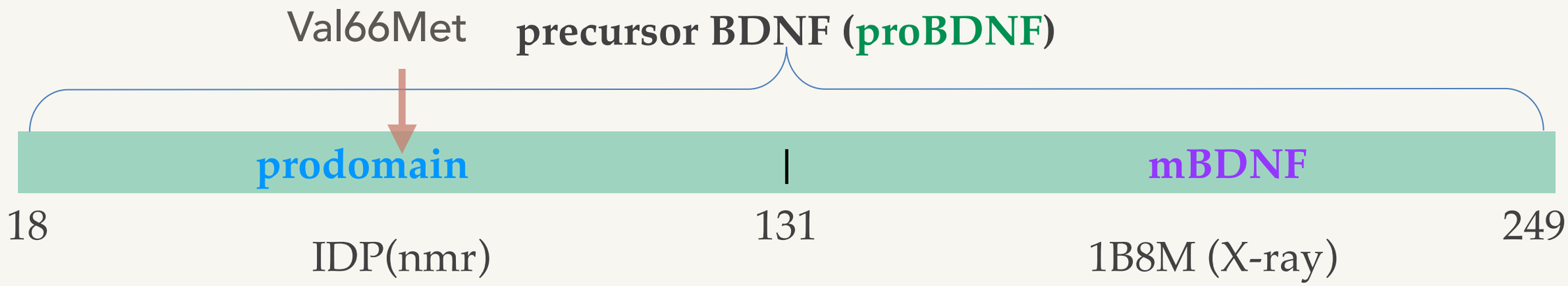
- Most of the mutational studies on IDPs deals with charged residues i.e either changing the composition or the pattern of the oppositely charged residues

Tertiary contacts

- Mutations in IDP's can cause secondary structure changes at residues far away from mutation cite, however, how does mutation modulates long range contacts is not well understood

(Ganguly et al.,PLoS Comput Bio, 2015)

Protein of interest: Brain-derived neurotrophic factor (BDNF) and Val66MET SNP



- Neurotrophin family of signaling proteins

- prodomain : Intracellular trafficking, folding of mature domain

- Val66Met SNP is most widely studied (>10,000 papers until 2017) and is associated with memory impairments among others

- NMR -> pro-domain is disordered with differential secondary structure preferences for V66 and M66

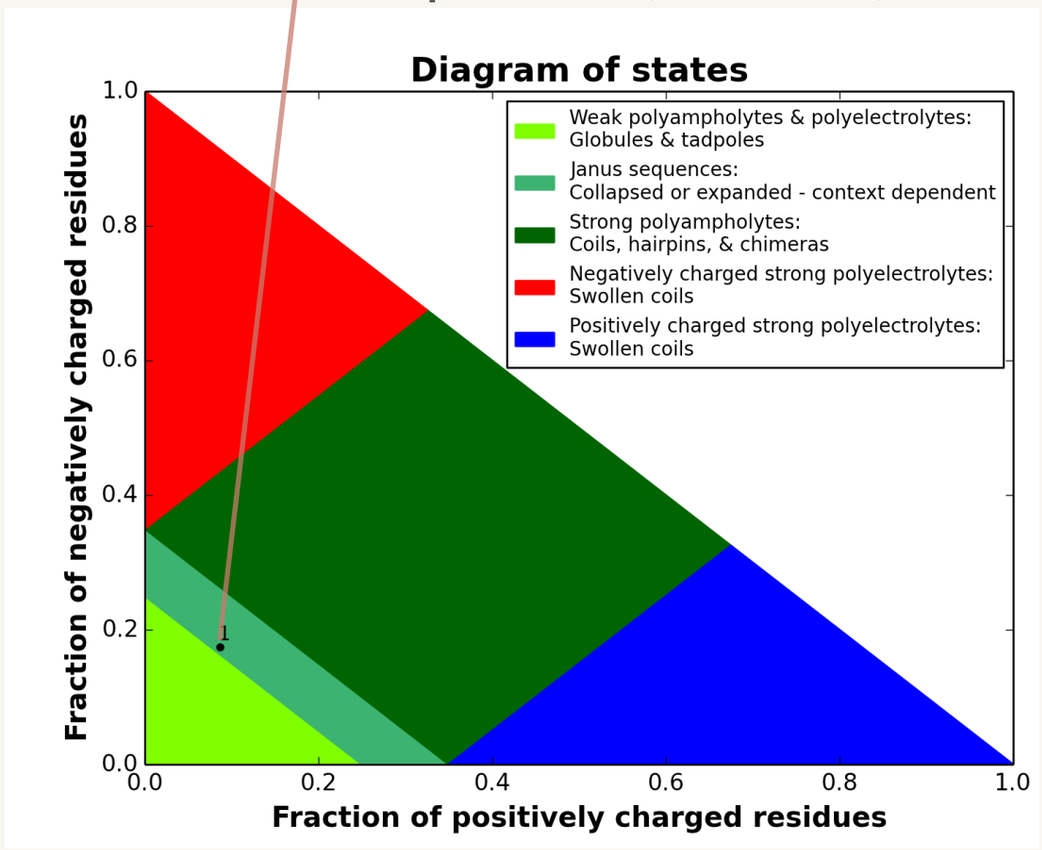
- Only Met66 proBDNF causes neuronal growth cone retraction by binding to SorCS2 (sortilin-related VPS10p-domain containing receptor 2)

(Anastasia et al.,Nat Commun., 2013)

Prediction of disorder based on prodomain sequence

- No charge difference between V66 and M66 sequence
- Both V66 and M66 is disordered using various disorder predictors

both V66 and M66 falls at the same point in janus region

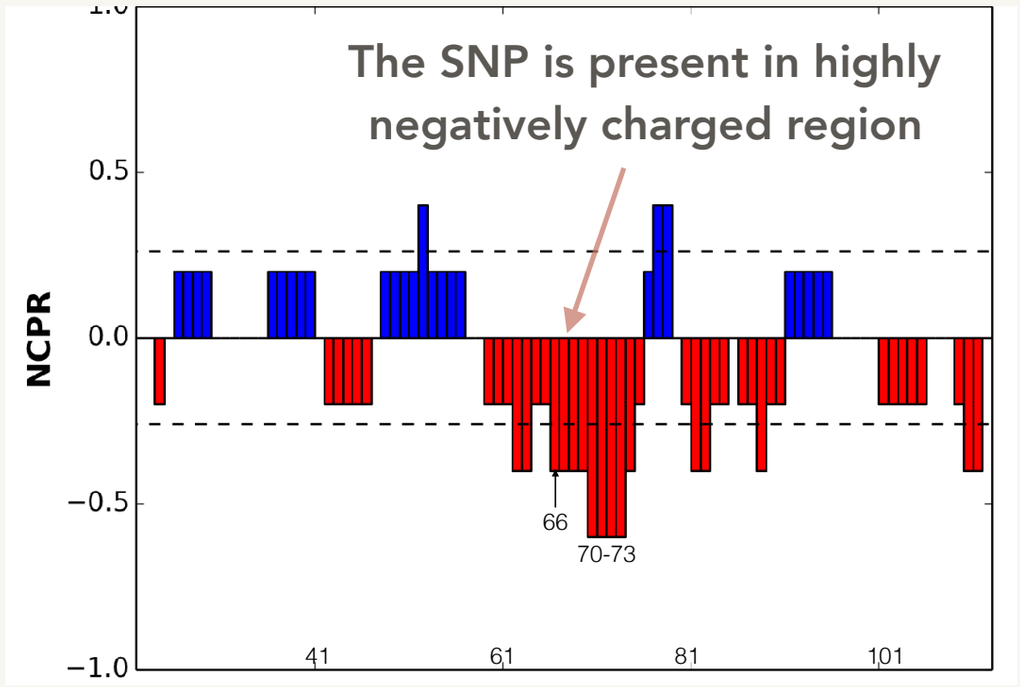


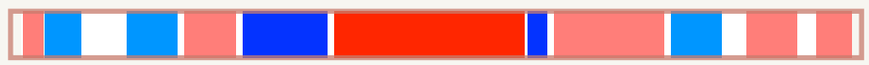
f_+ .09

f_- .17

κ .21

NCPR - net charge per residue based on sliding window of 5 residues





Main Questions

- Why does the V66M mutation affect residual local secondary structure?
- Does this mutation also affect the protein packing (Rg), even though it is charge-neutral?
- Is this effect mediated by changes in secondary structure, or by direct interactions of V vs M with rest of sequence?
- Is there a meaningful way to characterize tertiary structure of IDPs?

Method - MD Simulation

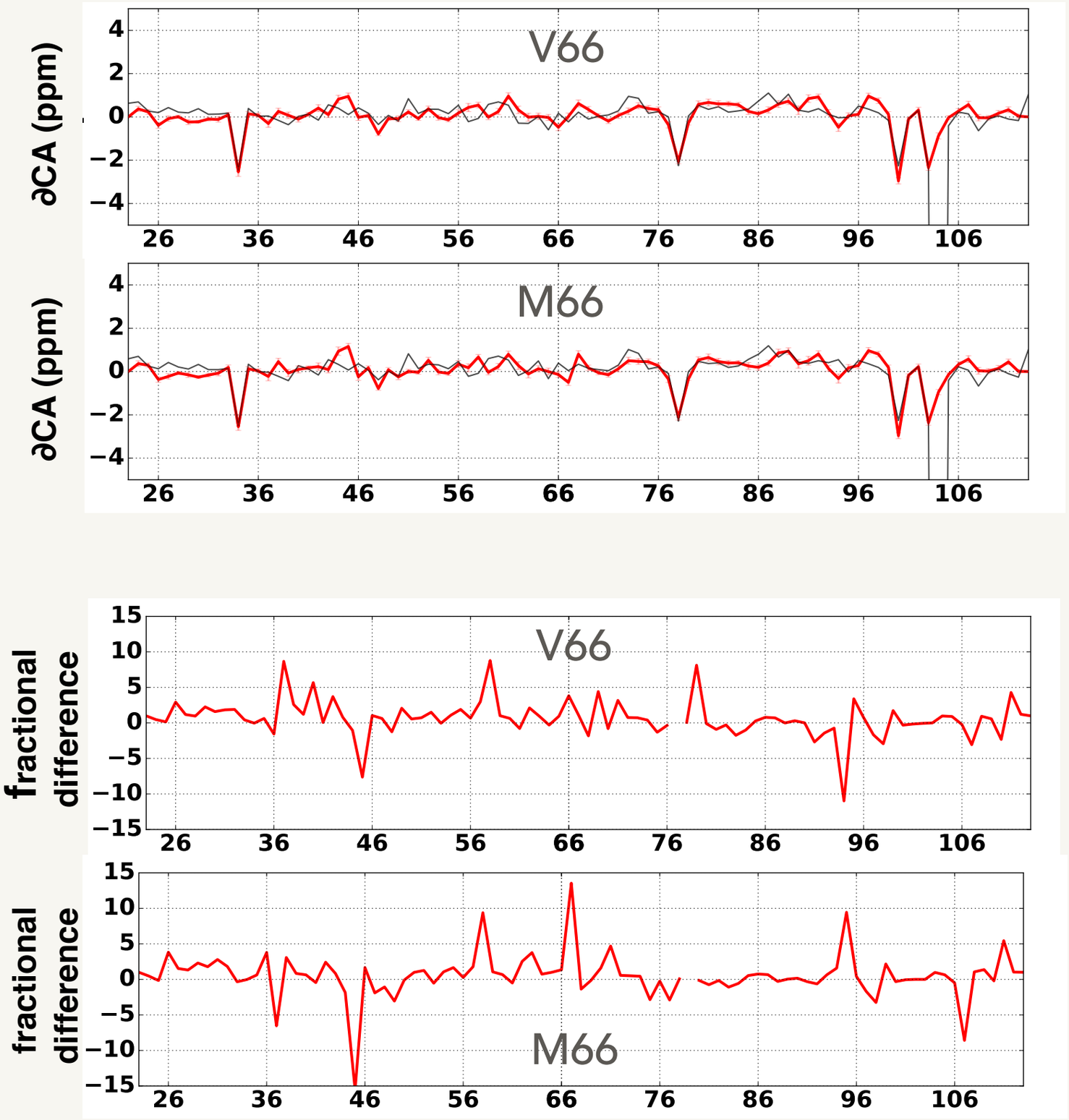
- The experimental method X-ray crystallography often fails for IDP's
- NMR gives us the average conformational ensemble properties
- Molecular dynamics (MD) simulations are an indispensable tool for studying IDP's. It gives us insight at microscopic levels
- 16μs explicit solvent replica exchange simulation of the prodomain region 23-113 (91 residues) for both Val and Met forms
- GROMACS 5.0.7 simulation package with Amber99sb-idln-q force field and TIP4p-D water model
- 64 replicas in the temperature range of **300K** to **385 K** were each run for 250 ns, exchanged every 1ps. Acceptance ratio was 20-25%
- Initial conditions : Different random coil for each replica (However, same starting structure for both V and M form)

Comparing with NMR data: CA chemical shift

NMR 273K

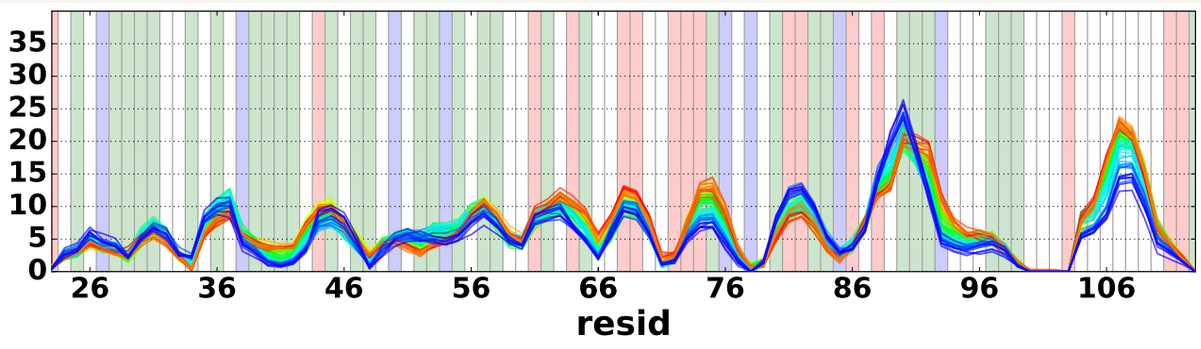
MD 300K

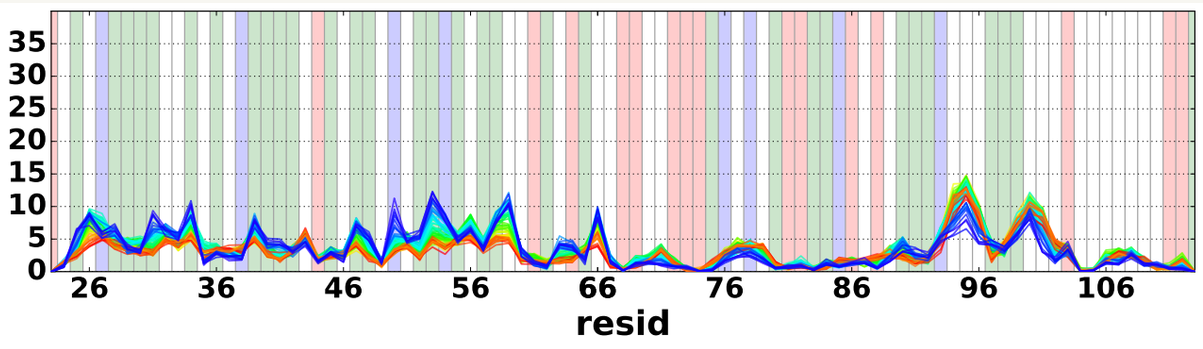
- Our simulations gives reasonable agreement with NMR



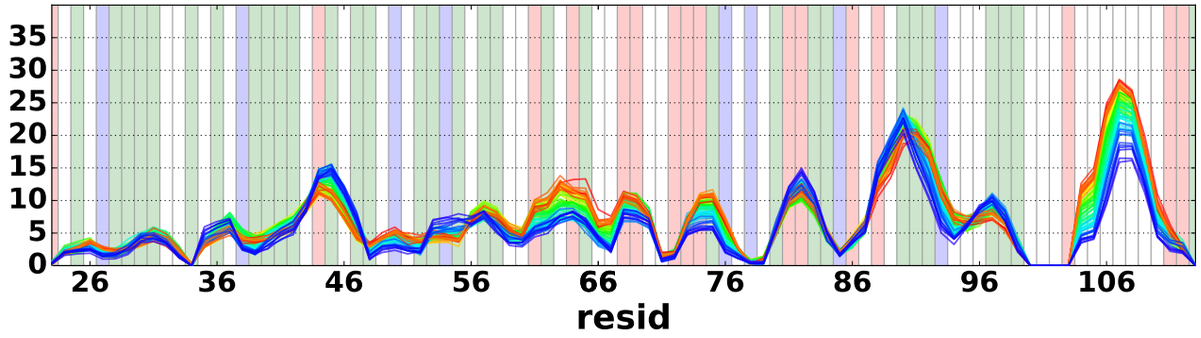
V66 vs M66 secondary structure

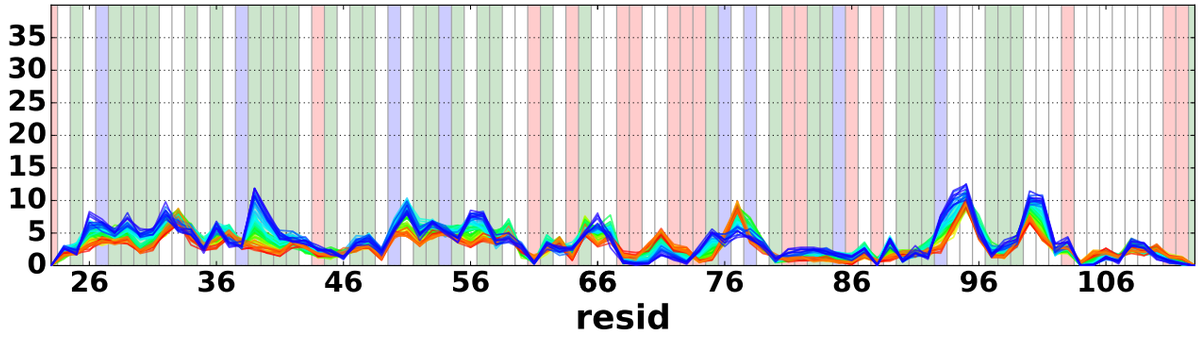
V66





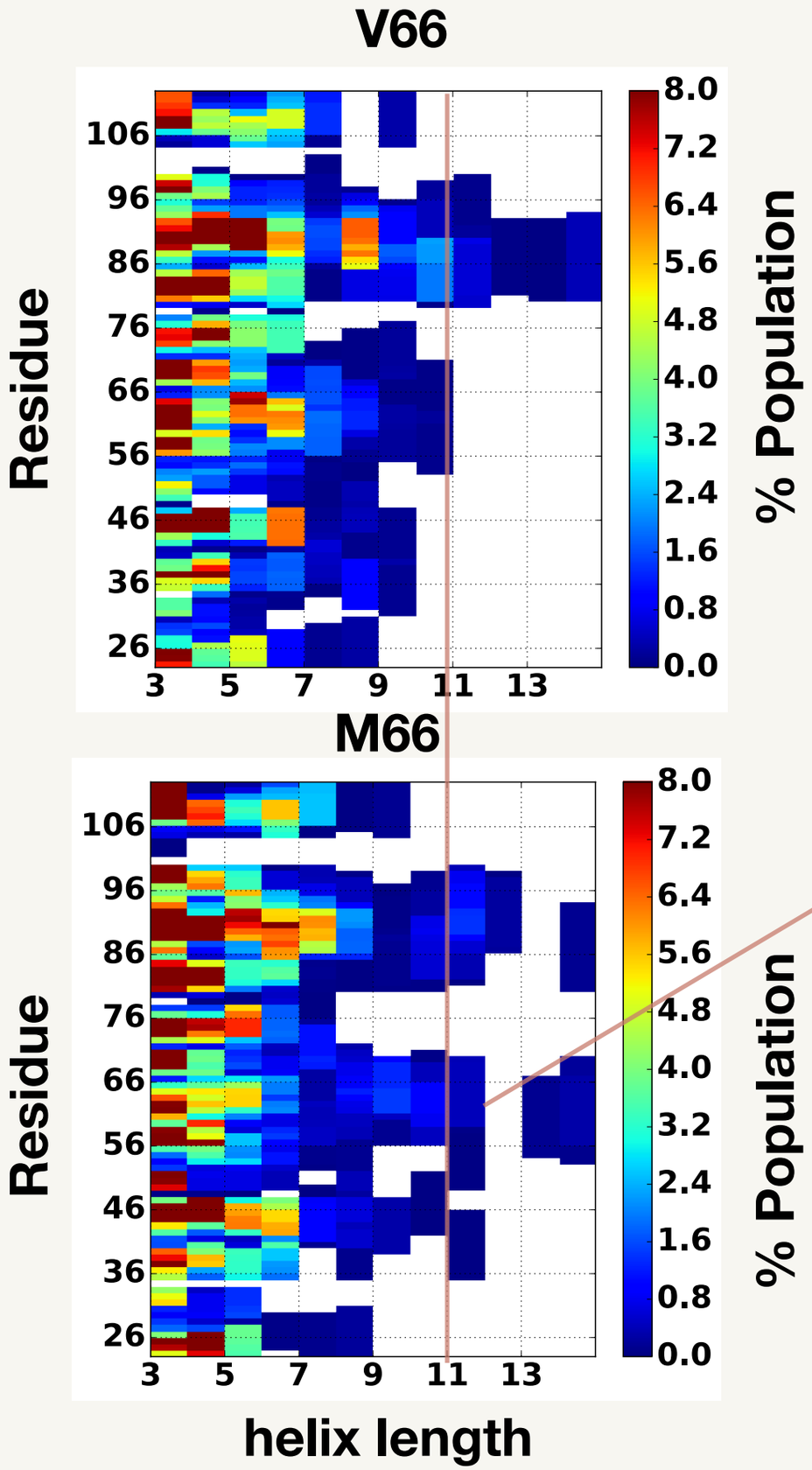
M66

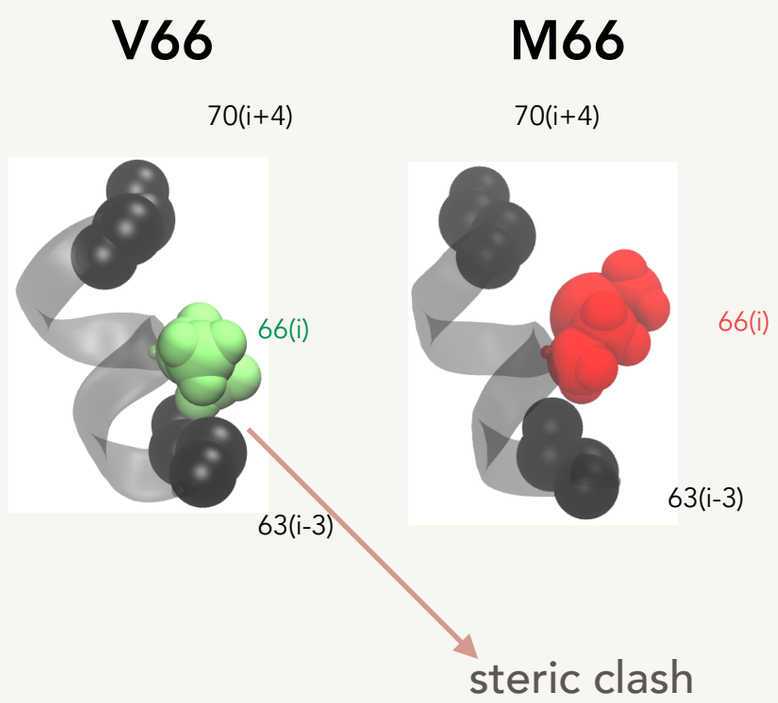




- Secondary structure assignment by STRIDE confirms disorder, with more secondary structure formation at higher temperatures

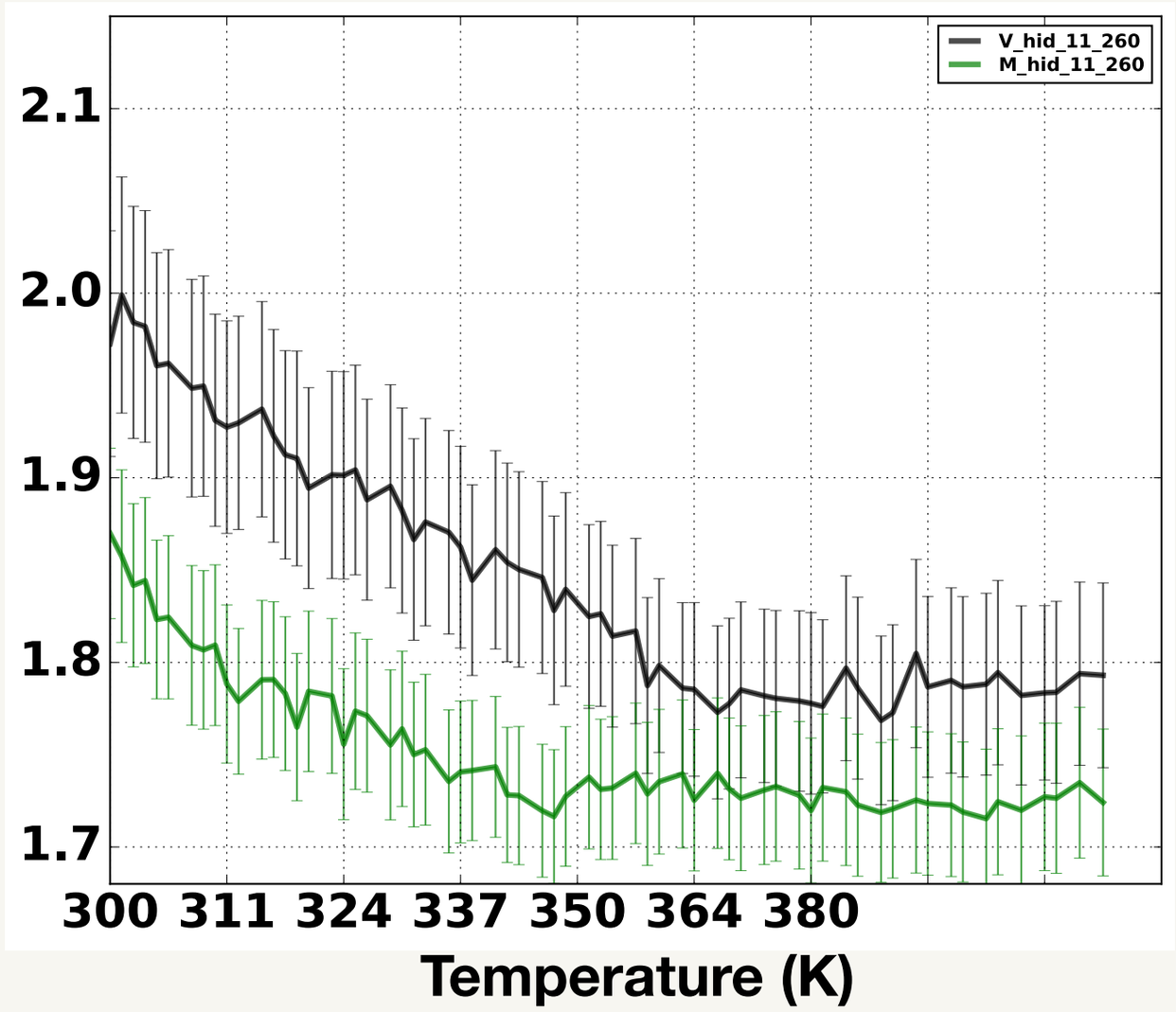
V66 vs M66 secondary structure : M66 forms longer helix at residue 66





- M66 forms longer helix at residue 66

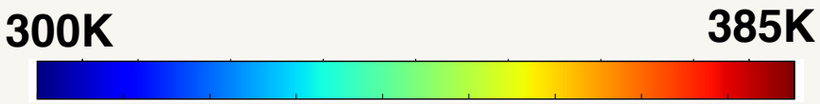
V66 vs M66 Radius of gyration (Rg)



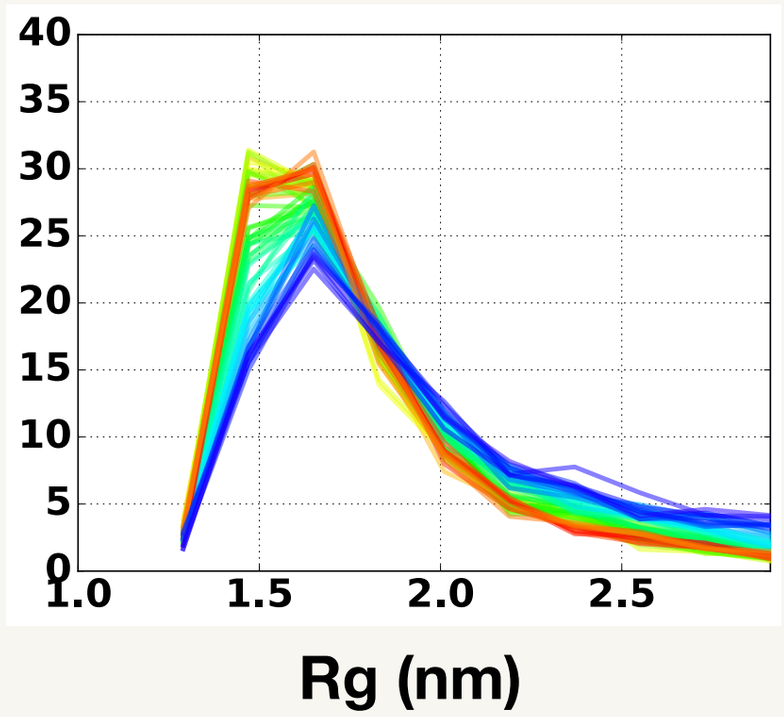
- NMR diffusion measurements reported slightly higher hydrodynamic radius for V66 than M66. (V66 - 2.24nm, M66 2.20nm at 298K).

- Simulations also indicate larger radius of gyration for V66 than M66 at 300K, and higher effective temperature for V66 compared to M66

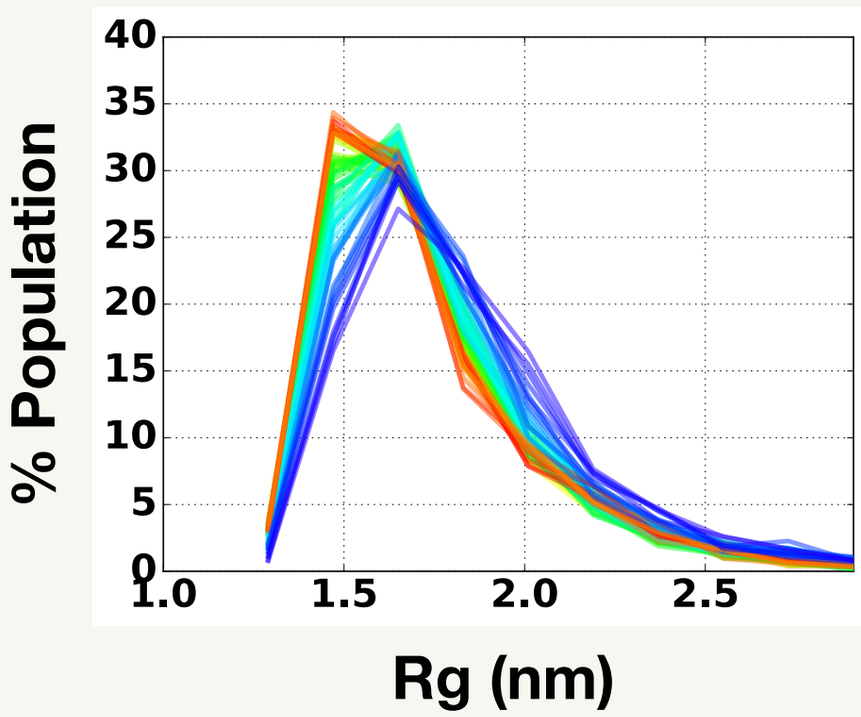
Radius of gyration (Rg): V66 has expanded conformations



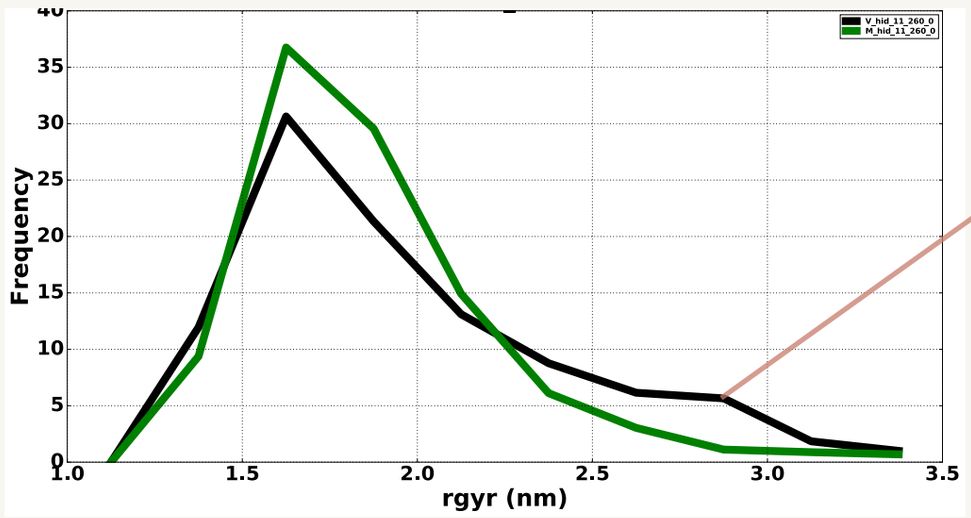
V66



M66



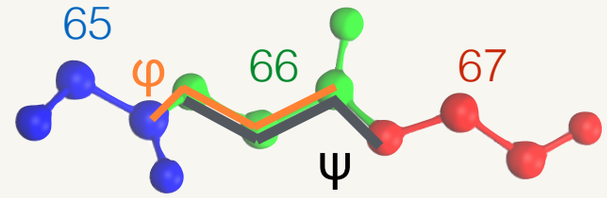
V66 vs M66 at 300K



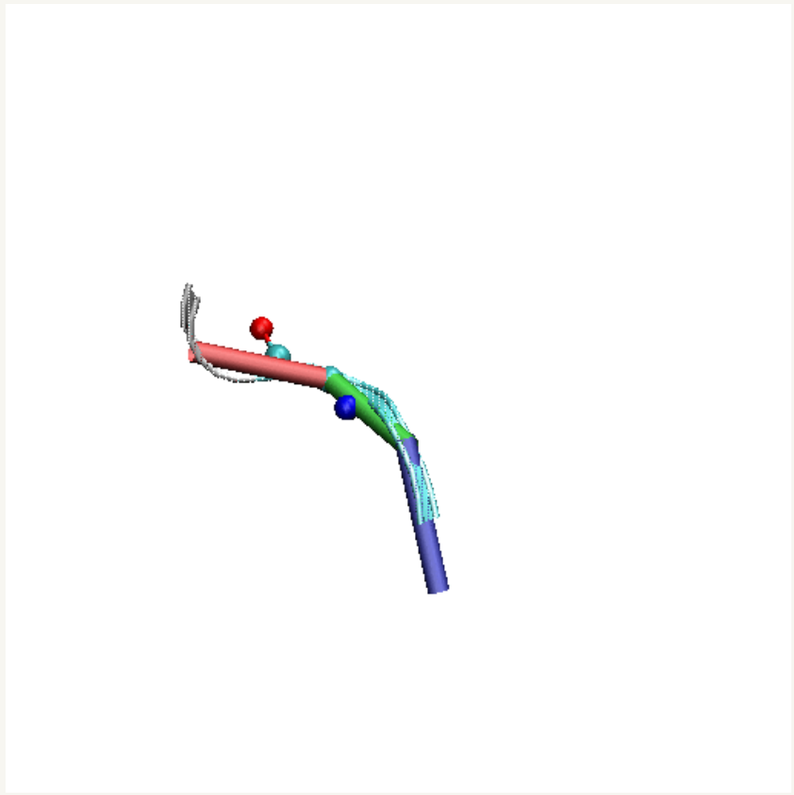
expanded structure

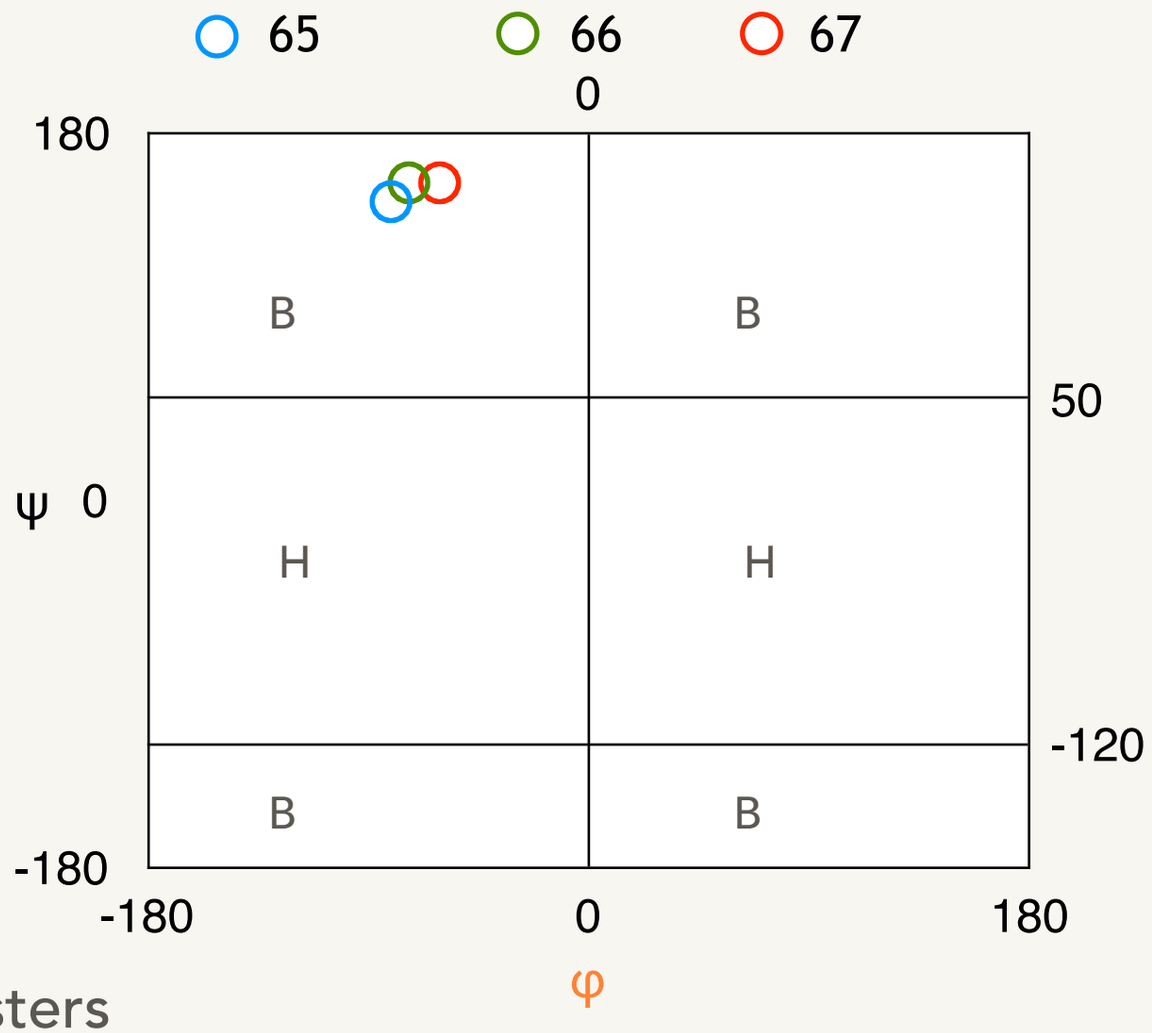
Clustering approach : dihedral angles for 65-66-67

$\phi = C_{i-1}\text{-}N\text{-}C_\alpha\text{-}C$
$\psi = N\text{-}C_\alpha\text{-}C\text{-}N_{i+1}$



- dihedral angles : orientation of a residue with flanking residues



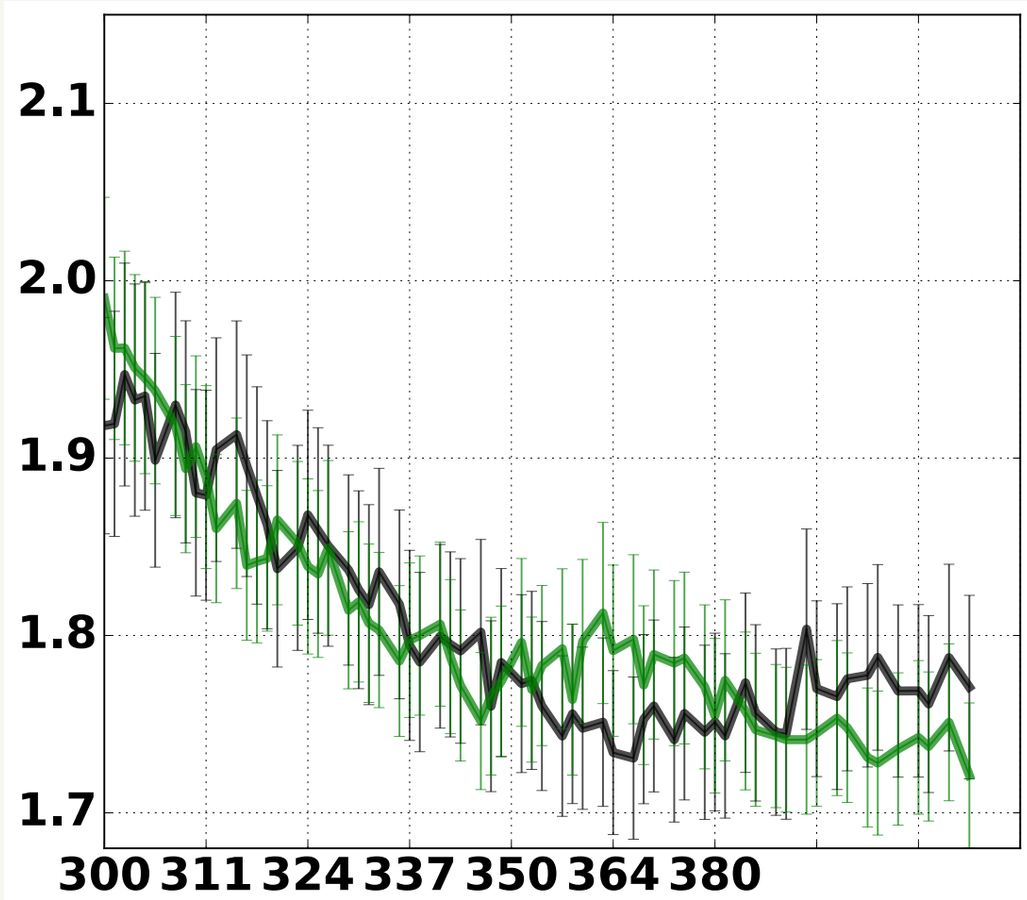


dihedral angles of 65,66 and 67 → 8clusters

HHH,BBB,BBH,HHB,HBB,BHH,BHB,HBH

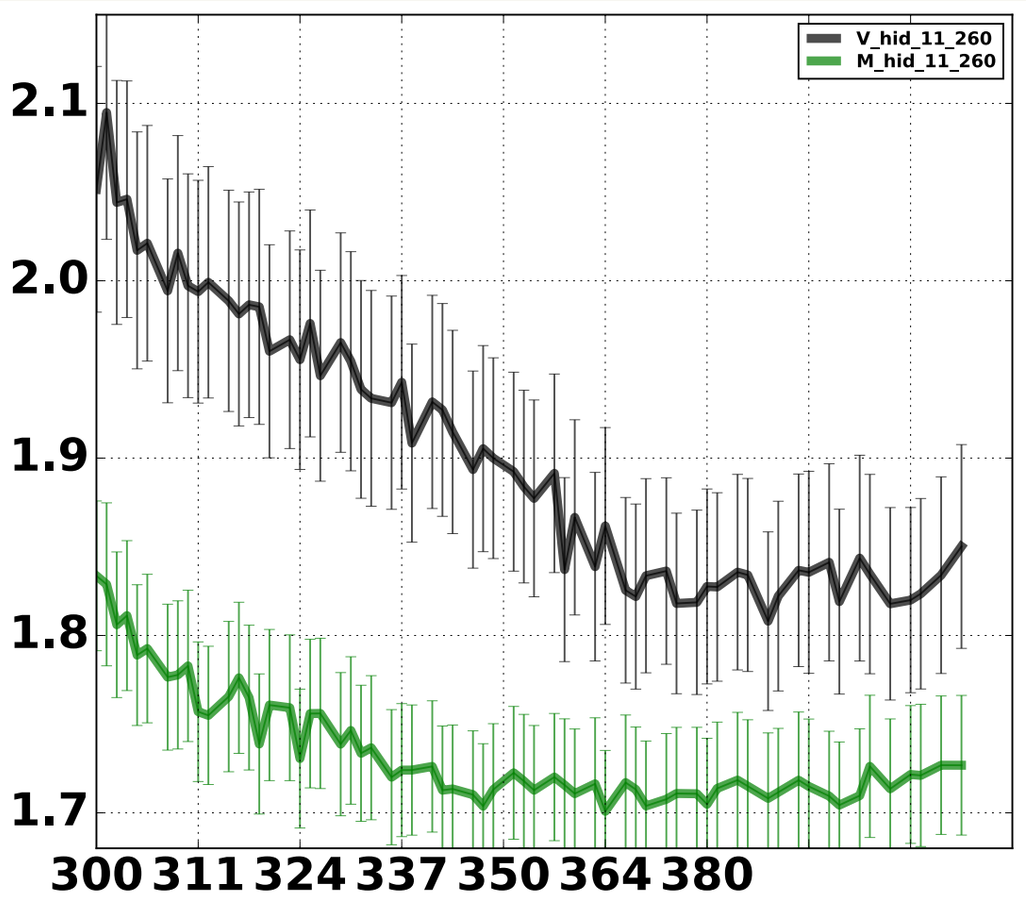
V66 vs M66 Radius of gyration (Rg)

Cluster-XXH



V66: 30%
M66: 23%

Cluster-XXB



V66: 70%
M66: 77%

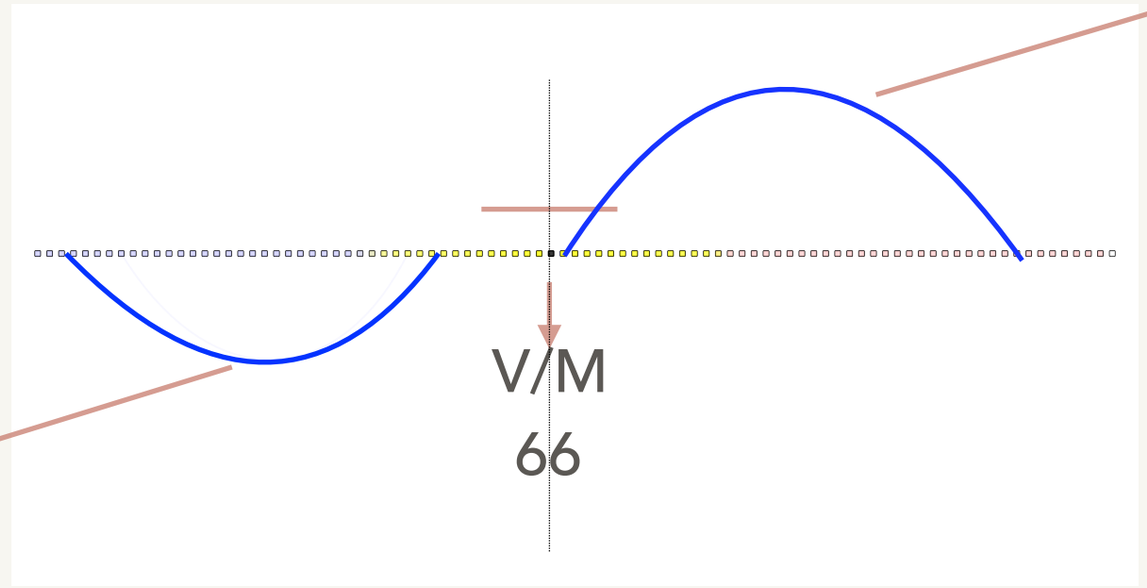
- Residue 67 in B conformation leads to high Rg

Pseudo-tertiary structure: long-range contacts at 300K

contacts present (top)

- If the contact formed involves residues 63 to 69

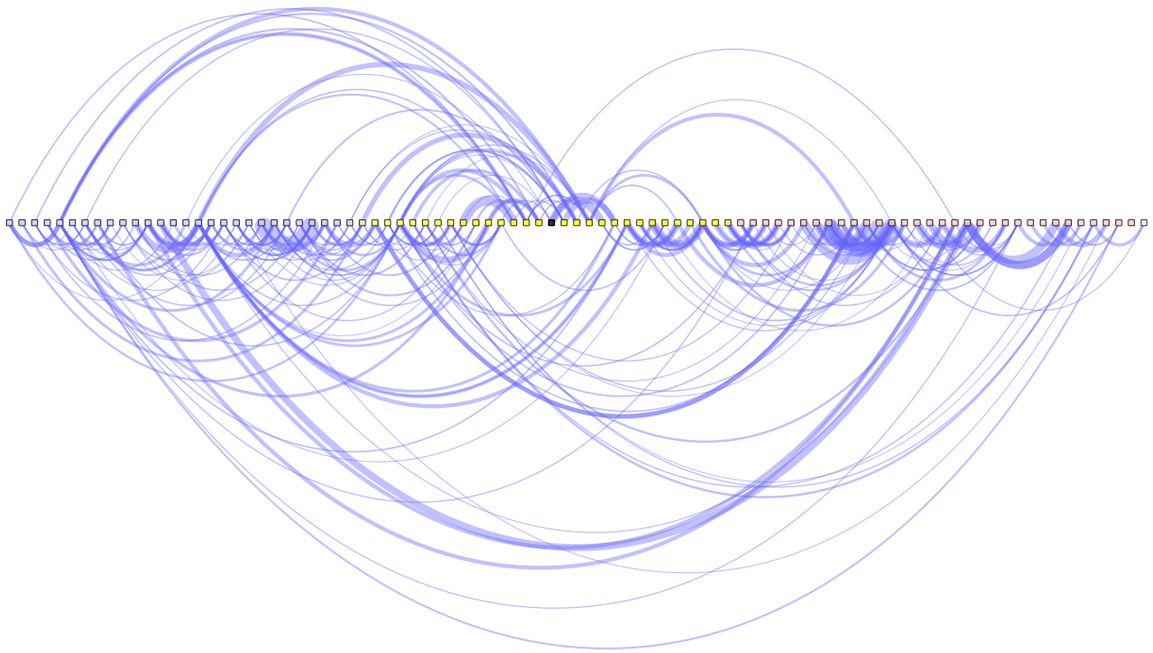
- The line thickness increases with increase in the difference between V66 and M66



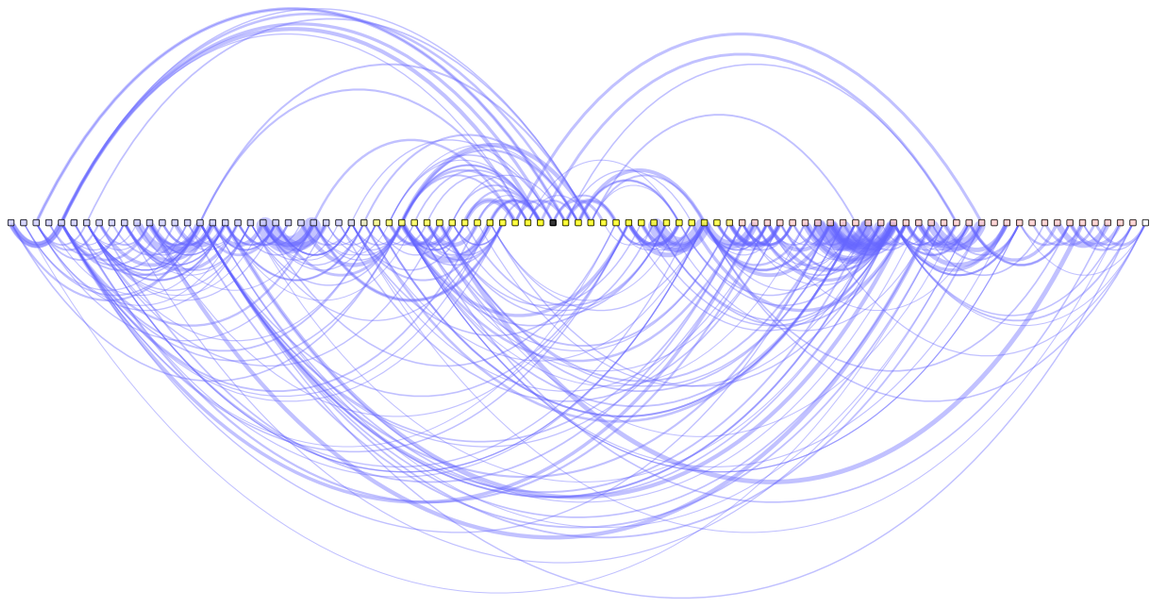
contacts present (bottom)

If the contact formed does not involves residues 63 to 69

V66



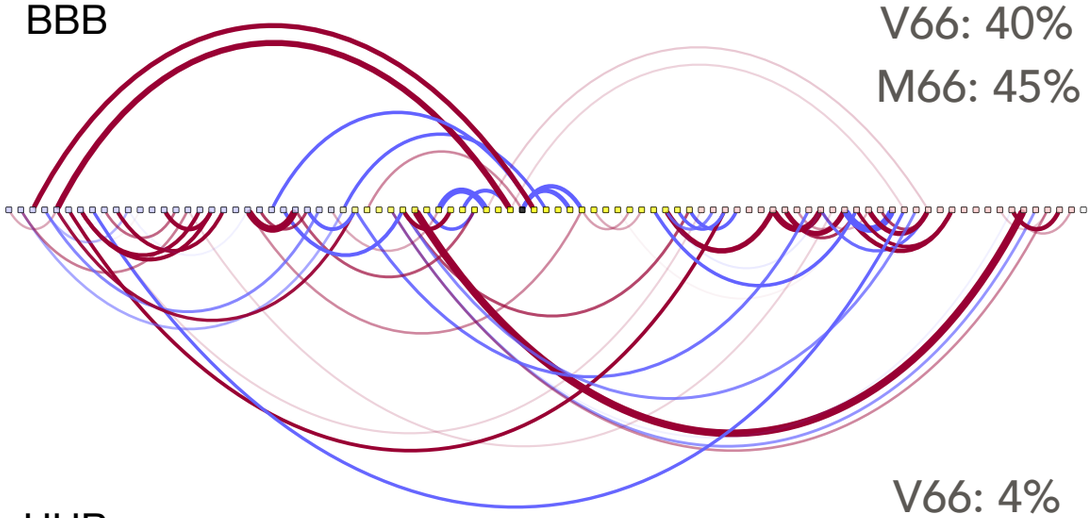
M66



V66 - M66, Cluster XXB

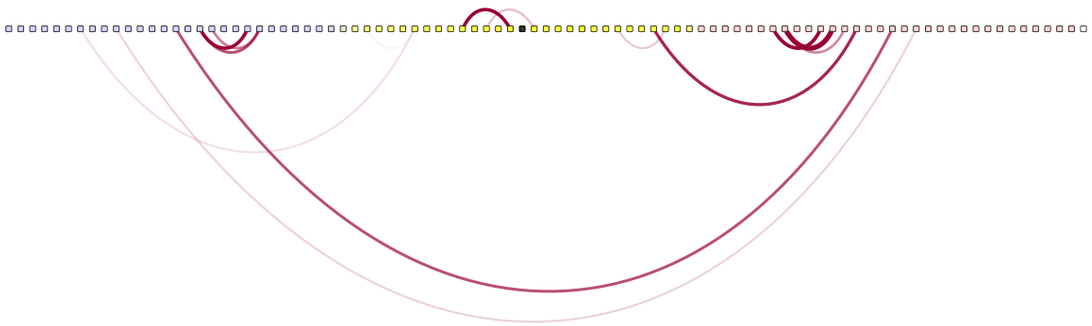
BBB

V66: 40%
M66: 45%



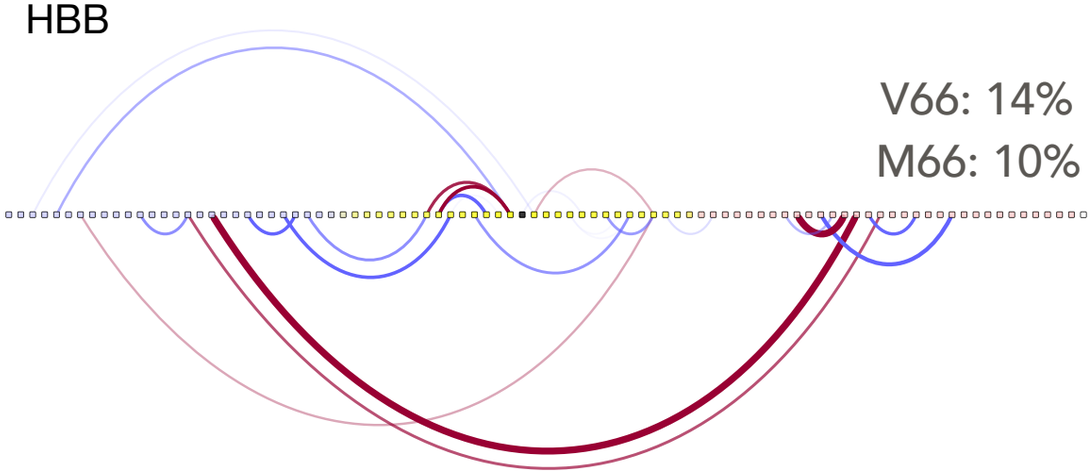
HHB

V66: 4%
M66: 12%



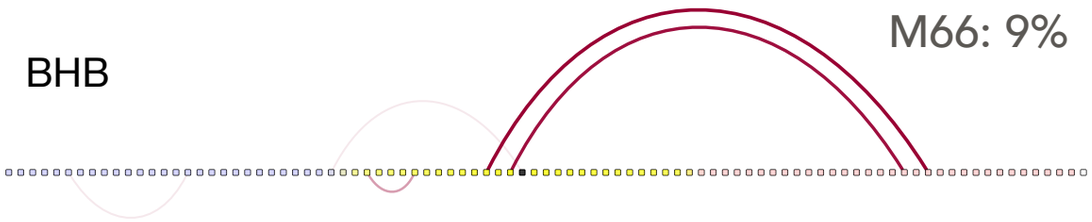
HBB

V66: 14%
M66: 10%



BHB

V66: 5%
M66: 9%



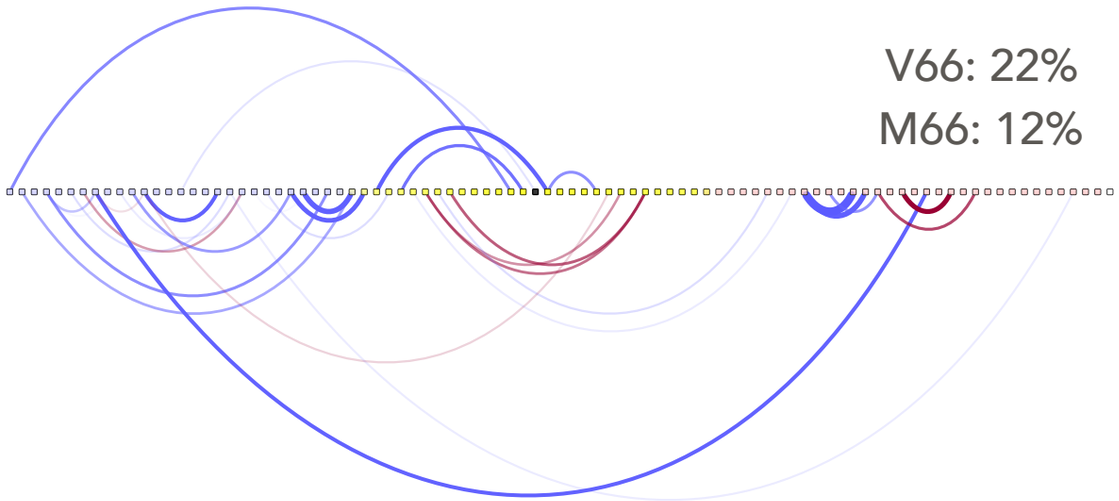
V66 - M66, Cluster XXH

HHH

V66: 6%
M66: 6%

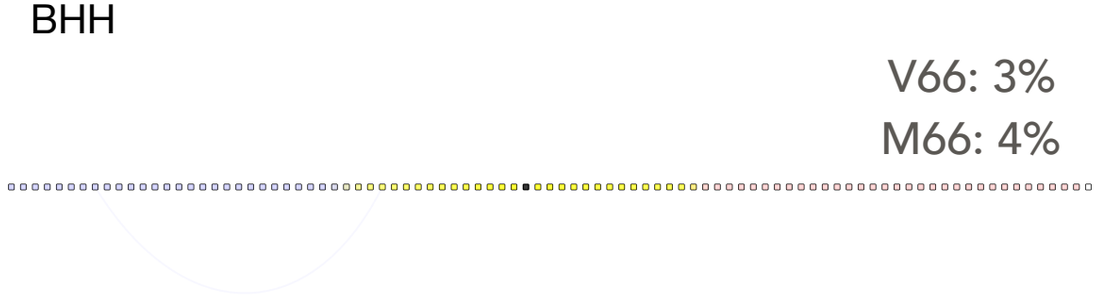
BBH

V66: 22%
M66: 12%



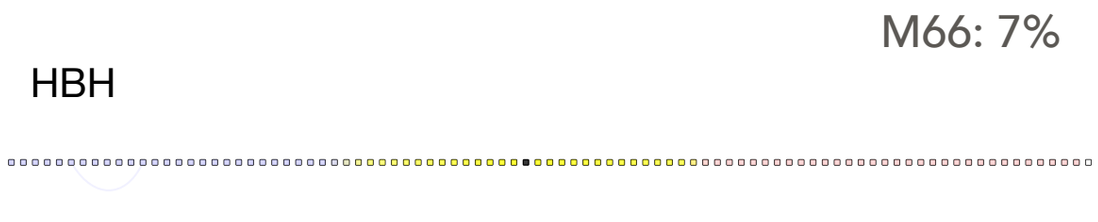
BHH

V66: 3%
M66: 4%

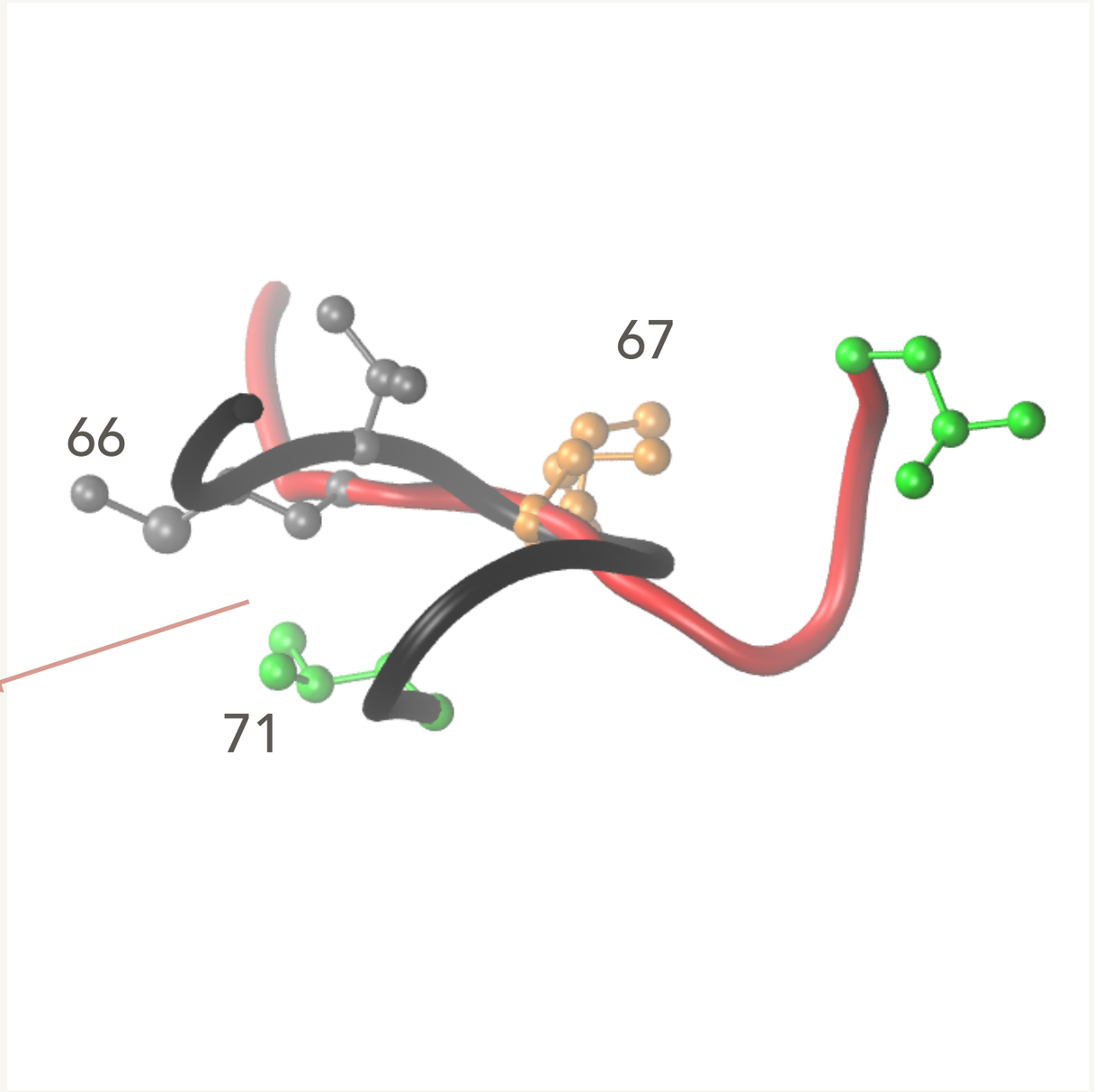


HBH

V66: 3%
M66: 7%



REPLACING RESIDUE 66 IN V66 WITH MET: GENERATES POTENTIAL CLASHES



- V66 forms short range contacts.
- Fewer long range contacts in V66 —> higher Rg
- 66-71 h-bond is entropically more favorable in V66 than in M66

Main Questions Underlying Research so far and their found answers

- Why does the V66M mutation affect residual local secondary structure? Differential entropic cost of helix formation
- Does this mutation also affect the protein packing (Rg), even though it is charge-neutral? Yes
- Is this effect mediated by changes in secondary structure, or by direct interactions of V vs M with rest of sequence? 66 affects secondary structure at 67 —> predicts tertiary contacts.
- Is there a meaningful way to characterize tertiary structure of IDPs? Yes

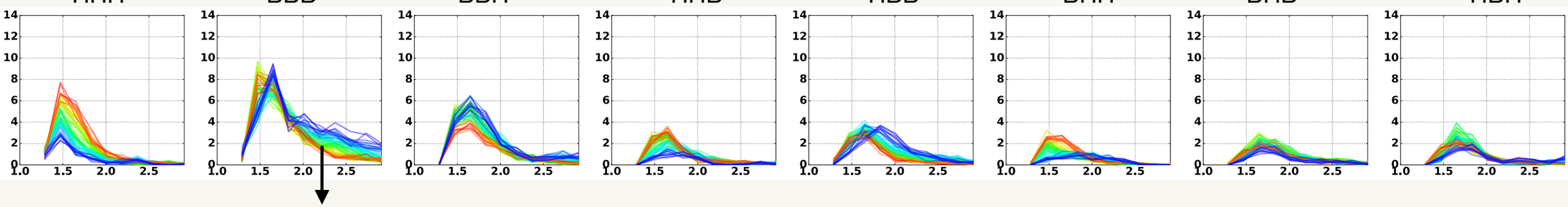
Acknowledgements

XSEDE, Stampede
Brannigan research group
My committee members:
Dr. Eric Klein
Dr. Jinglin Fu
Dr. Cameron Abrams
Dr. Grace Brannigan

Thank you for listening!

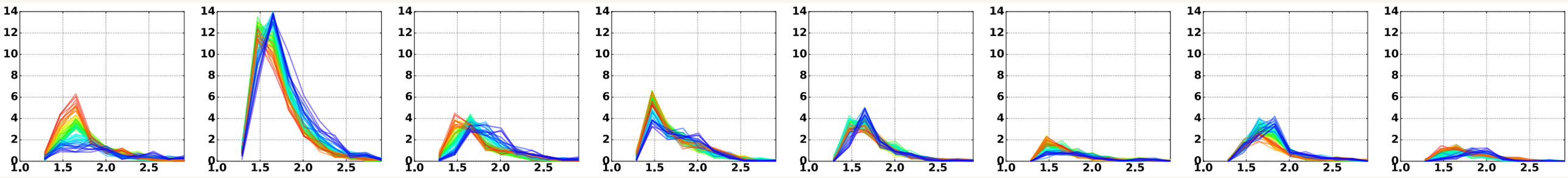
Rg distribution for every cluster

V66



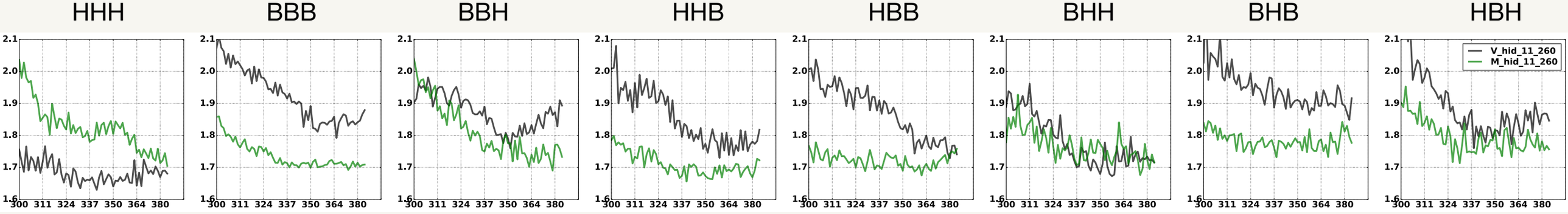
High Rg is observed in cluster BBB

M66



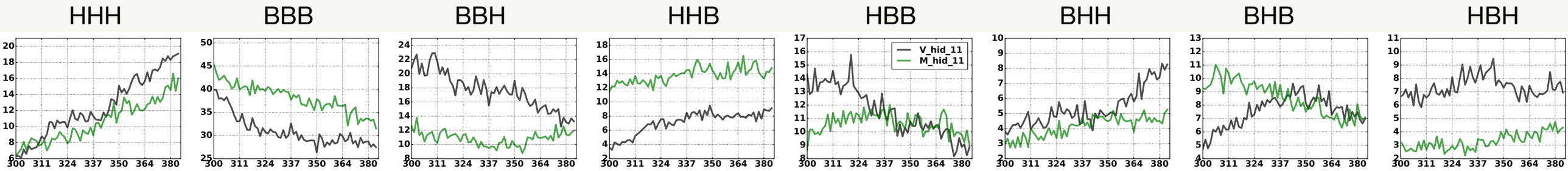
Rg (nm)

<Rg> for each cluster



The ensemble is broken down into 8 clusters depending on the arrangement of residue 65, 66 and 67

%population of 8 clusters with temperature



Old- simulations