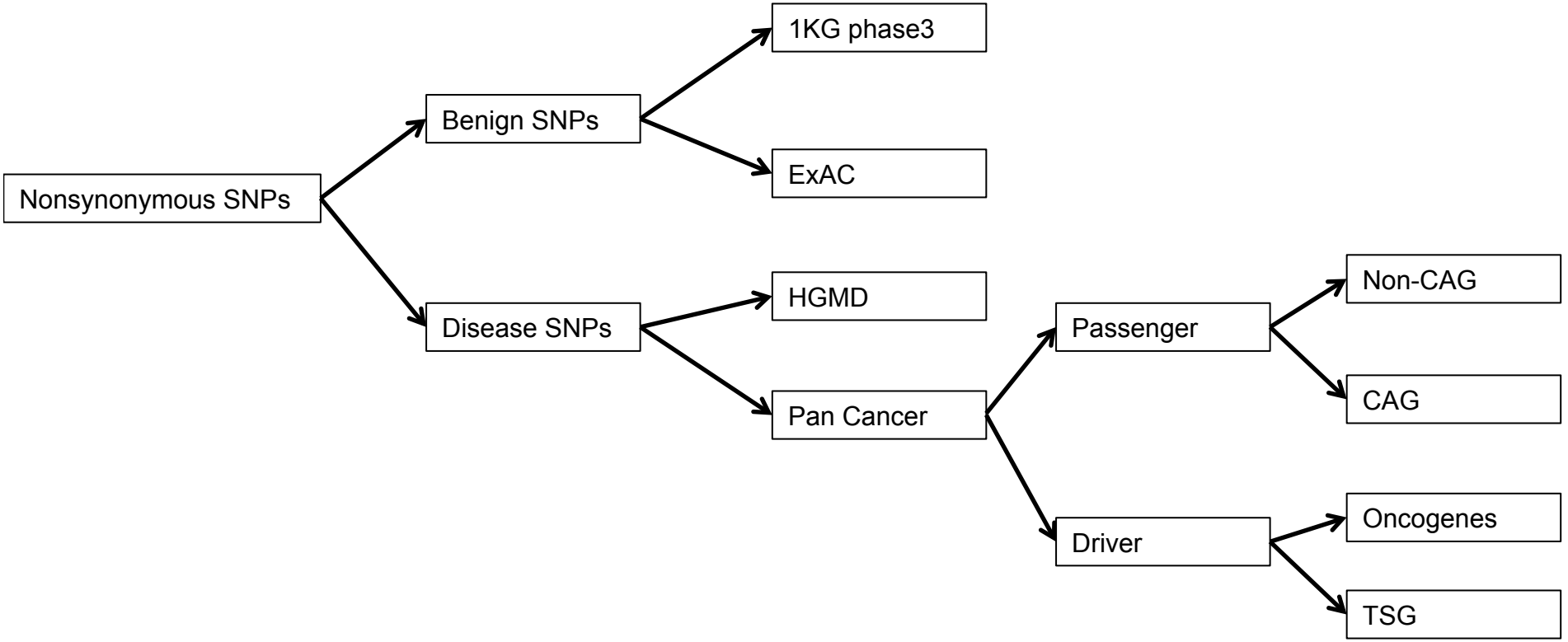
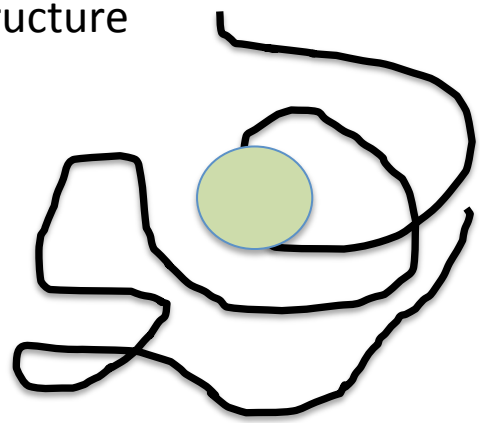
Frustration Analysis

SNV datasets for frustration analysis



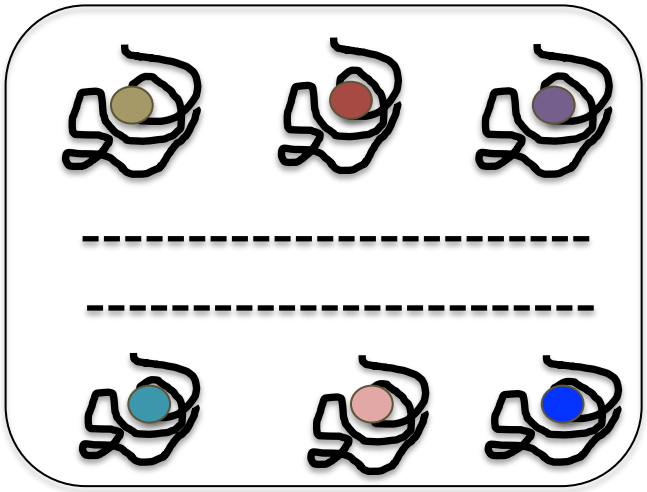
Native structure



$$F_i = \frac{E_i^{T,N} - \langle E_{i'}^{T,U} \rangle}{\sqrt{1/N \sum_{k=1}^{n} (E_{i'}^{T,U} - \langle E_{i'}^{T,U} \rangle)^2}}.$$

F_i >= 0.78 (minimal frustrated)
F_i <= -1.0 (maximal frustrated)

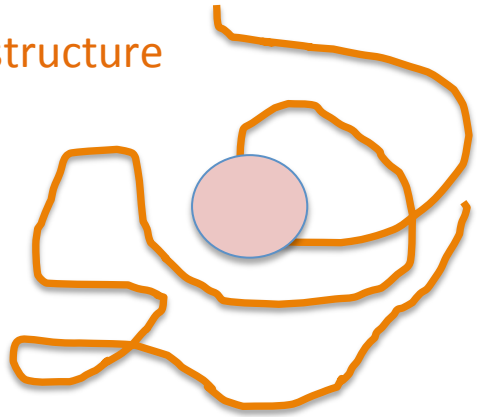
Native Residue frustration Index (NRFI)



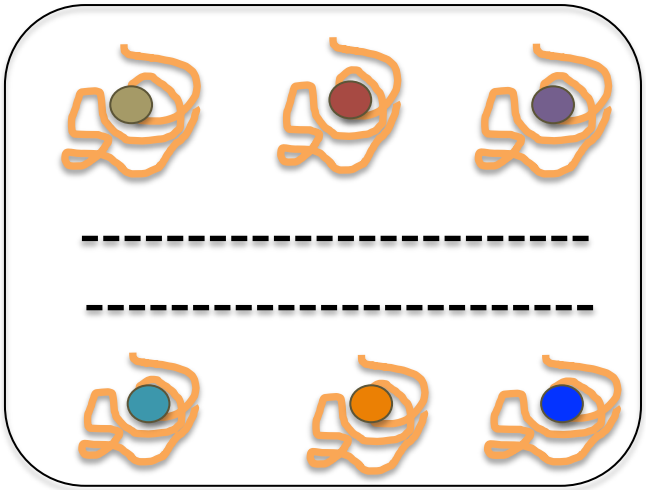
HOMOLOGY MODELLING

Δ Residue frustration = MRFI - NRFI

Mutated structure



Mutated Residue frustration Index (MRFI)



Maximal frustration

large amount of conflicting interactions, local interactions not satisfied.

example -> hydrophilic residues in the core of folded protein

Minimal frustration

no conflicting interactions, all local interactions satisfied.

example -> hydrophobic residues in the core of folded protein

Delta frustration = Mutated Residue frustration – Native residue frustration

For maximally frustrated residues in the native state:

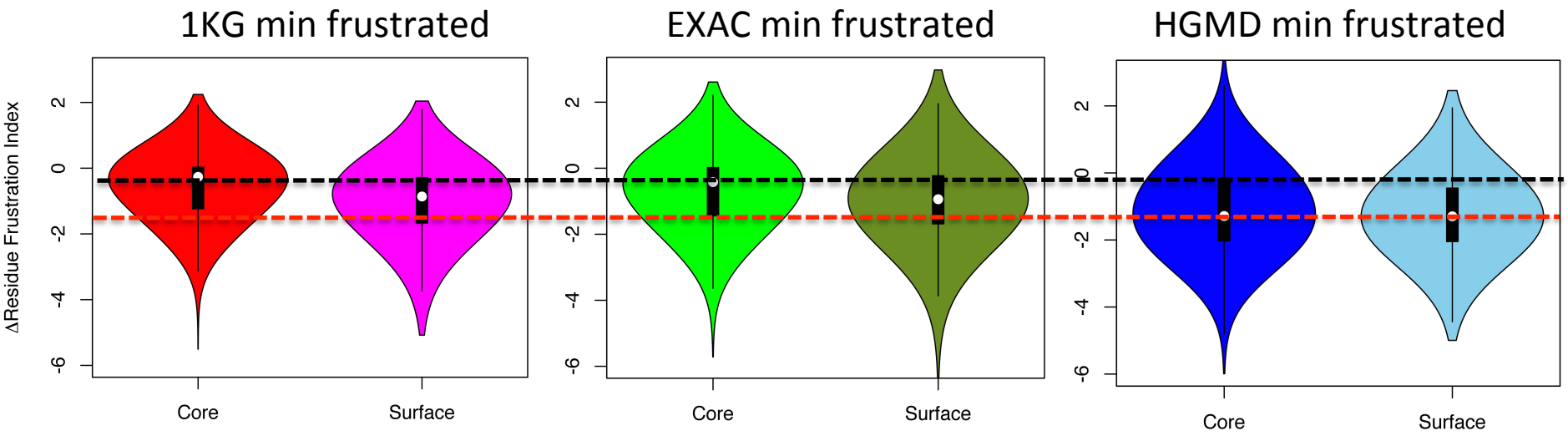
After Mutation

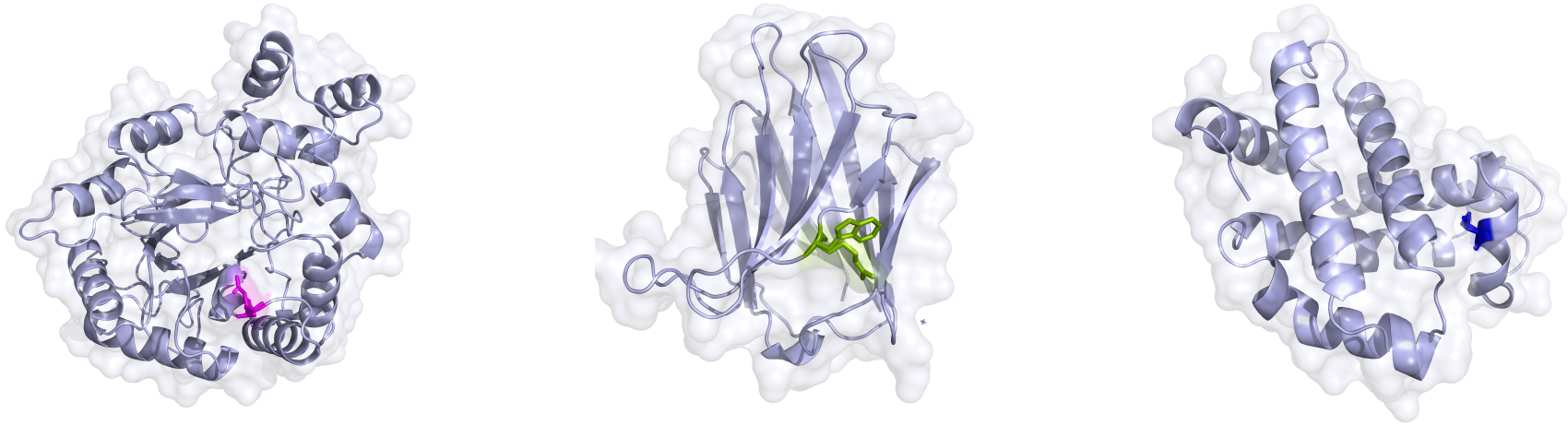
Delta frustration > 0 (expected most of the time) -> release (decrease in conflicting interactions) of frustration

For minimally frustrated residues in the native state:

After Mutation

Delta frustration < 0 (expected most of the time) -> gain (increase in conflicting interactions) of frustration

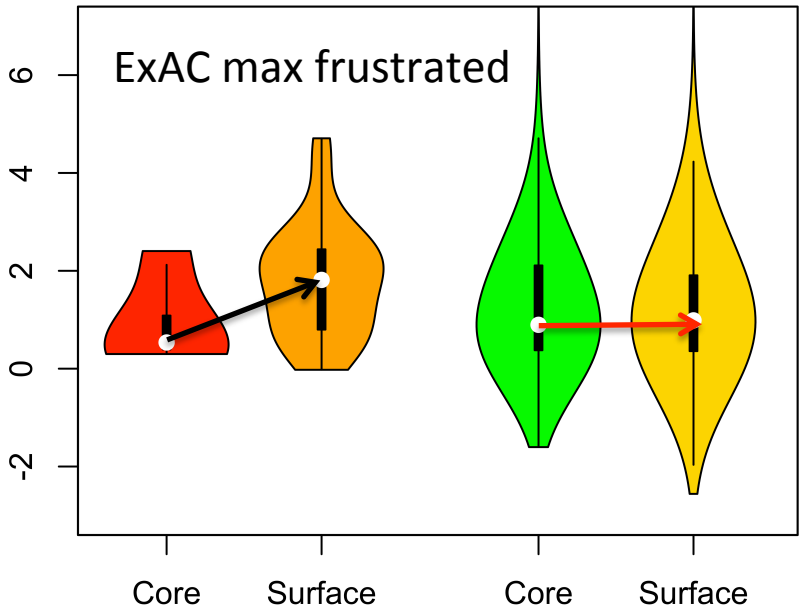


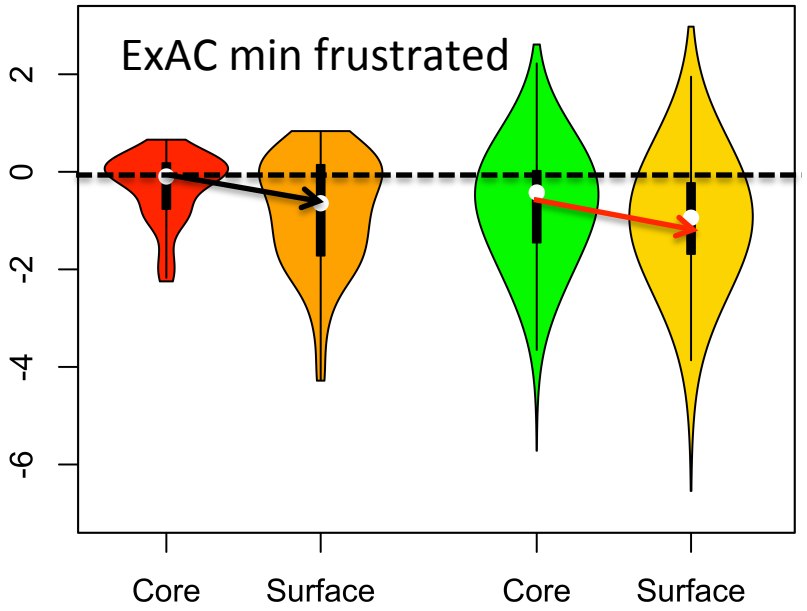


Disease mutations (HGMD) lead to higher frustration gain (larger delta negative) compared to benign mutations (1KG,EXAC). **Disease mutations disrupt the localized stability of protein residues to a greater extent**.

Benign mutations (1KG & EXAC) lead to higher frustration gain in surface compared to core residues.

However, average **frustration gain is approximately same for core and surface residues impacted by disease mutations**.





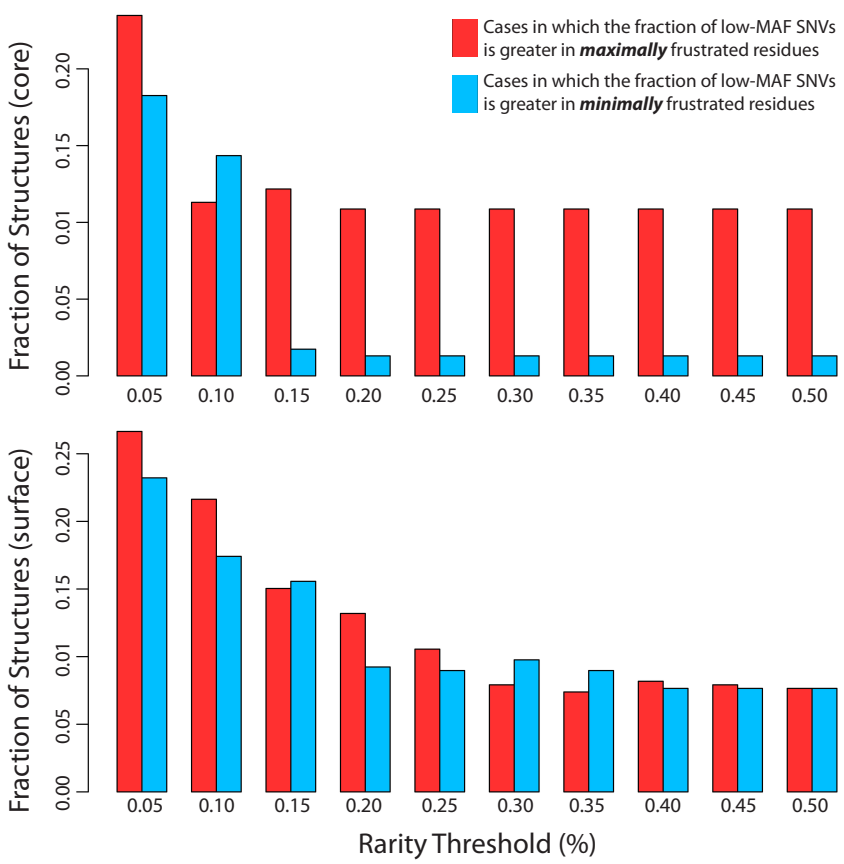
Common Core
Common Surface
Rare Core
Rare Surface

For maximally frustrated residues, common variants influencing surface residues tend to lose higher amount of frustration compared to core residues.

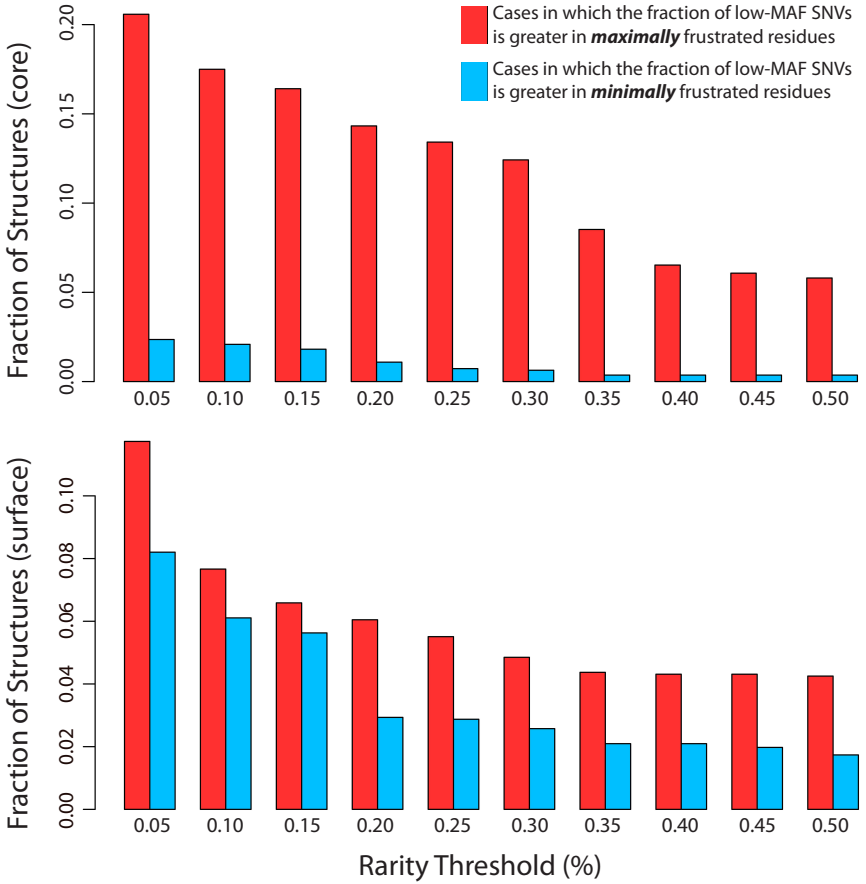
For minimally frustrated residues , presence of rare variants lead to higher gain in frustration for both surface and core residues. **Higher disruptive effect of rare variant compared to common variants on localized interactions of a given residue.**

For minimally frustrated residues, surface residues gain more frustration compared to core residues. **Surface residue interactions getting more disrupted compared to core residue.**

1KG SNVs



ExAC SNVs

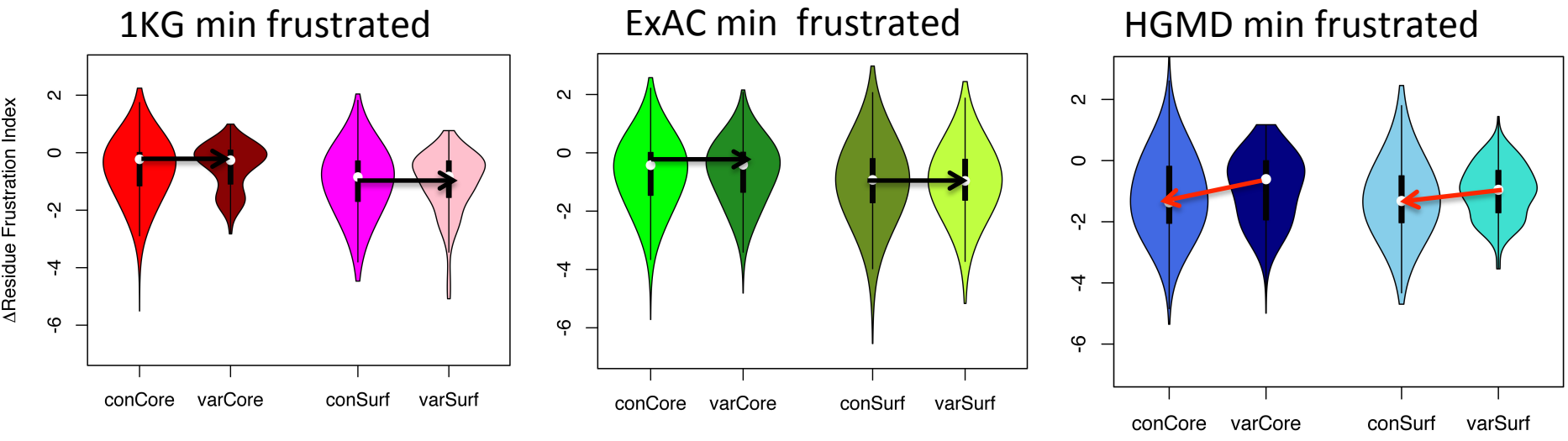


Using the fraction of rare variants to quantify negative selection, maximally frustrated residues tend to be more conserved than minimally frustrated residues.

Biologically, this could be a result of the functional roles played by maximally frustrated residues (allostery, etc)?

The trend is more pronounced in the core.

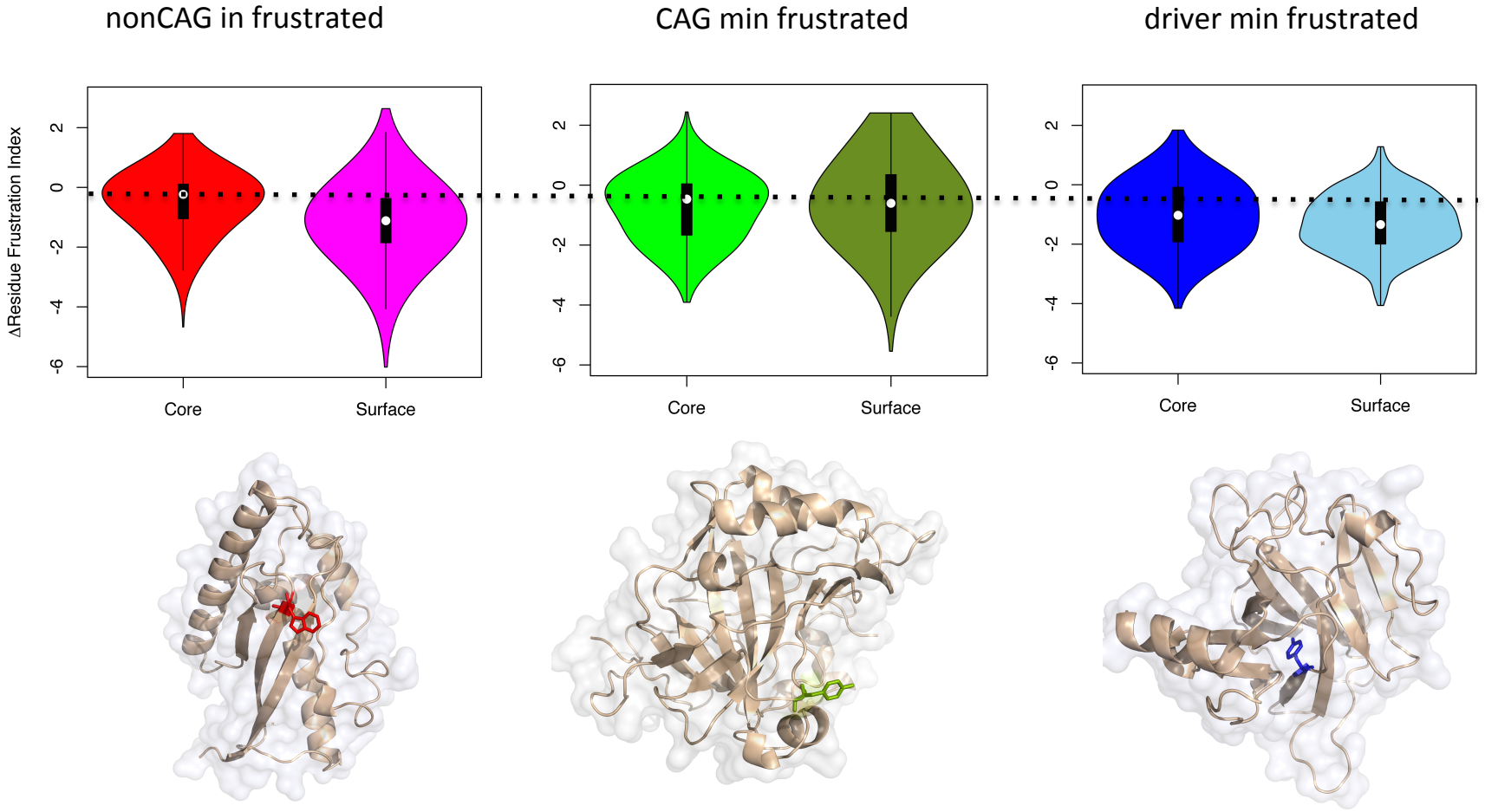
GERP score based analysis



No significant difference in delta frustration between benign mutations originating in conserved and variable region of the genome.

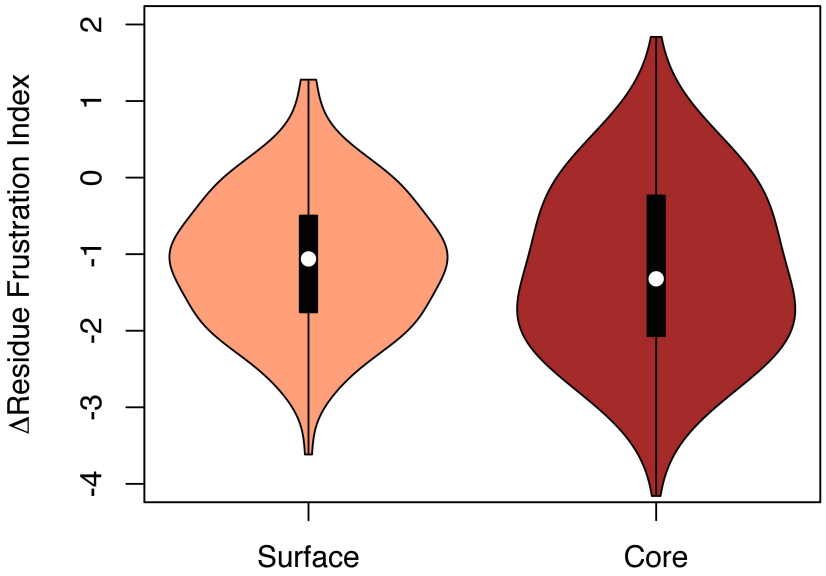
Disease mutations affecting conserved region of the genome lead to higher gain in frustration compared to mutations fixated in the variable region of the genome.

This difference is more pronounced in core residues compared to surface residues.

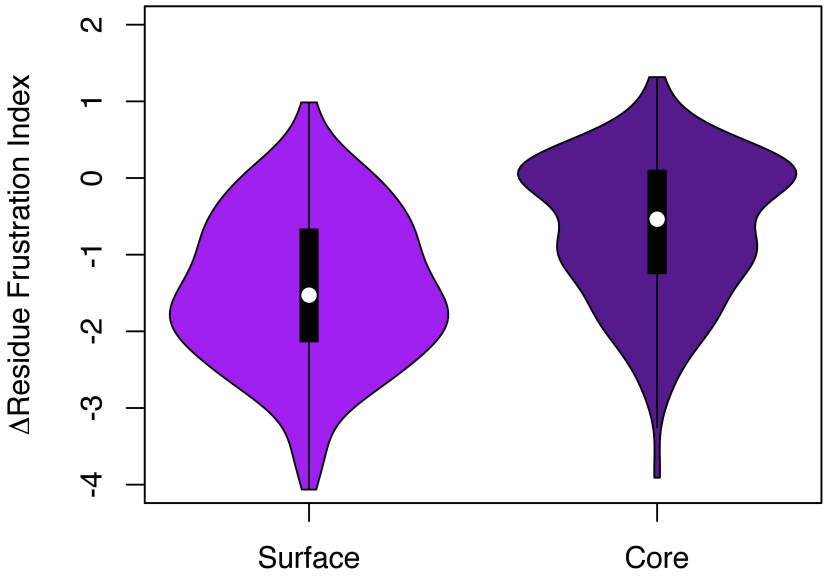


Overall cancer driver mutations gain more frustrations while affecting minimally frustrated residues.

There is no significant difference in gain of frustrations for passenger mutations impacting cancer associated genes (CAG) and otherwise (non-CAG).



TSG min frustrated



Oncogene min frustrated

Driver mutation affecting tumor suppressor genes, which maps onto core residues, gain more frustration compared to surface residues.

Oncogene driver mutations mapping onto surface residues gain significant amount of frustration compared to core residues.

The above observation suggest two distinct mechanism by which driver mutations can disrupt local stability profile of residues in a protein.

A) TSG influencing driver mutation prefer disrupting the core residues interaction leading to destabilization of the hydrophobic core

B) Whereas **driver mutations affecting Oncogene disrupt surface residues stability**
Hypothesis -> in order to alleviate this higher disruption in surface residues, Oncogene encoded proteins might interact with other proteins non-specifically.