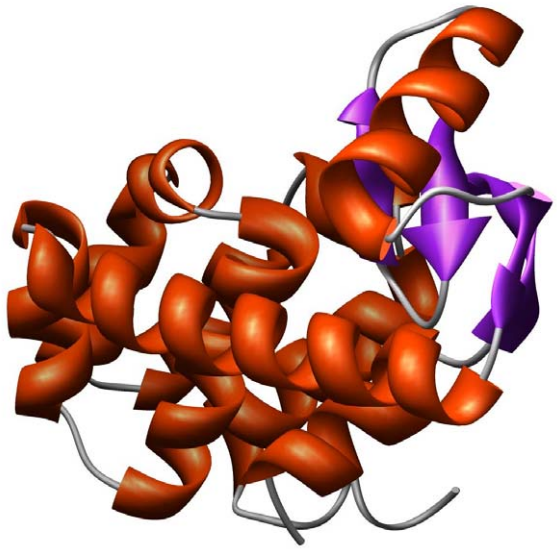
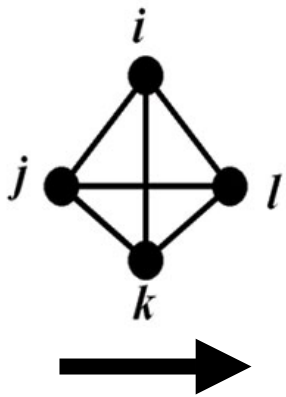
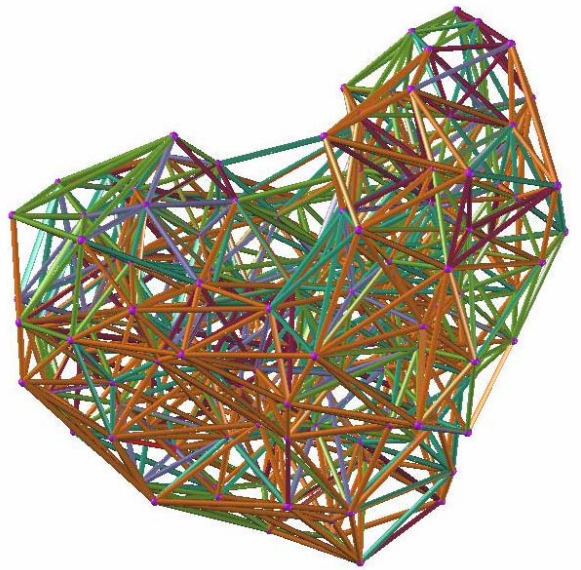
A Structure-Based Computational Mutagenesis Elucidates the Spectrum of Stability-Activity Relationships in Proteins







Majid Masso (mmasso@gmu.edu)

Laboratory for Structural Bioinformatics, School of Systems Biology

George Mason University, Manassas, Virginia

Slides available for download at http://binf.gmu.edu/mmasso

Previous Work and Motivation

- Auto-Mute (http://proteins.gmu.edu/automute): tools to predict protein stability and activity changes upon single residue mutation
- No simple relationship between these measures
  - Separate models for each prediction task
  - Each model trained using diverse mutant dataset with already known activity or stability change levels
- Cannot infer one property based on knowledge of the other
  - Evidence for "stability-function" hypothesis (increased stability at cost of activity, in some enzyme active sites); inverse trade-offs more controversial
  - Mutants for which both properties change in the same direction considered anomalous and generally ignored in the literature
- Here we attempt a comprehensive study of mutations, located throughout diverse proteins, with known values for both properties

Structure-Based Representation of Mutants

- Makes use of a computational mutagenesis methodology that we previously developed
- Yields a measure that we refer to as the mutant “Residual Score”
- The mutant residual score quantifies the relative change in protein sequence-structure compatibility due to the substitution
- A collection of mutants belonging to the same category can be characterized by their mean residual score (MRS)

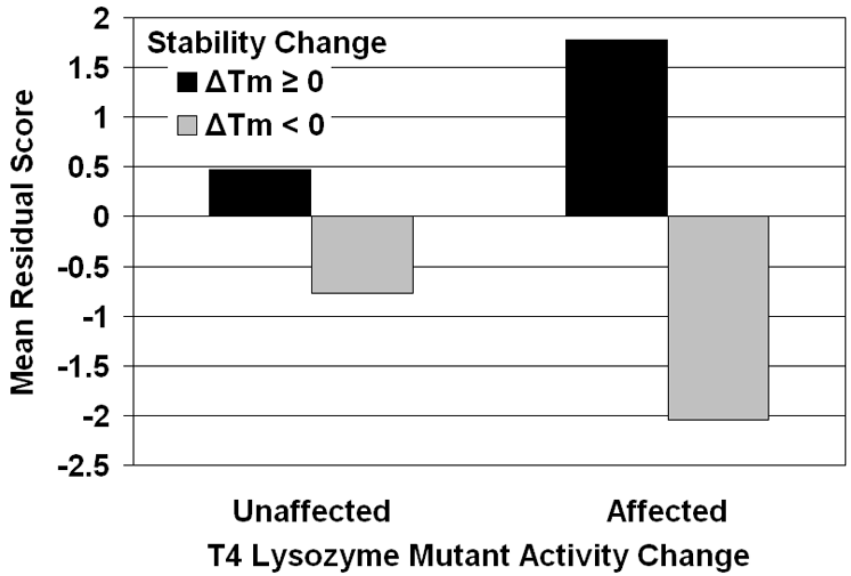
Motivating Example: Phage T4 lysozyme

- Rennell *et al*. synthesized and qualitatively measured activity of 2015 (65%) mutants, each labeled unaffected (U) or affected (A)
  - We previously calculated residual scores, and mean residual score (MRS) of the mutants in each activity class: MRS(U) = –0.76, MRS(A) = –1.40
  - Statistically significant structure (MRS) – function (U/A activity categories) relationship in T4 lysozyme (*t*-test, $p < 0.001$)

- Saraboji *et al.* studied 171 T4 lysozyme mutants with previously published stability change (ΔTm) values
  - Mutants collected from ProTherm database (repository of published data)
  - 121 of these mutants overlap with Rennell *et al.*
  - Categorize as increasing (inc, $\Delta Tm \geq 0$) or decreasing (dec, $\Delta Tm < 0$)

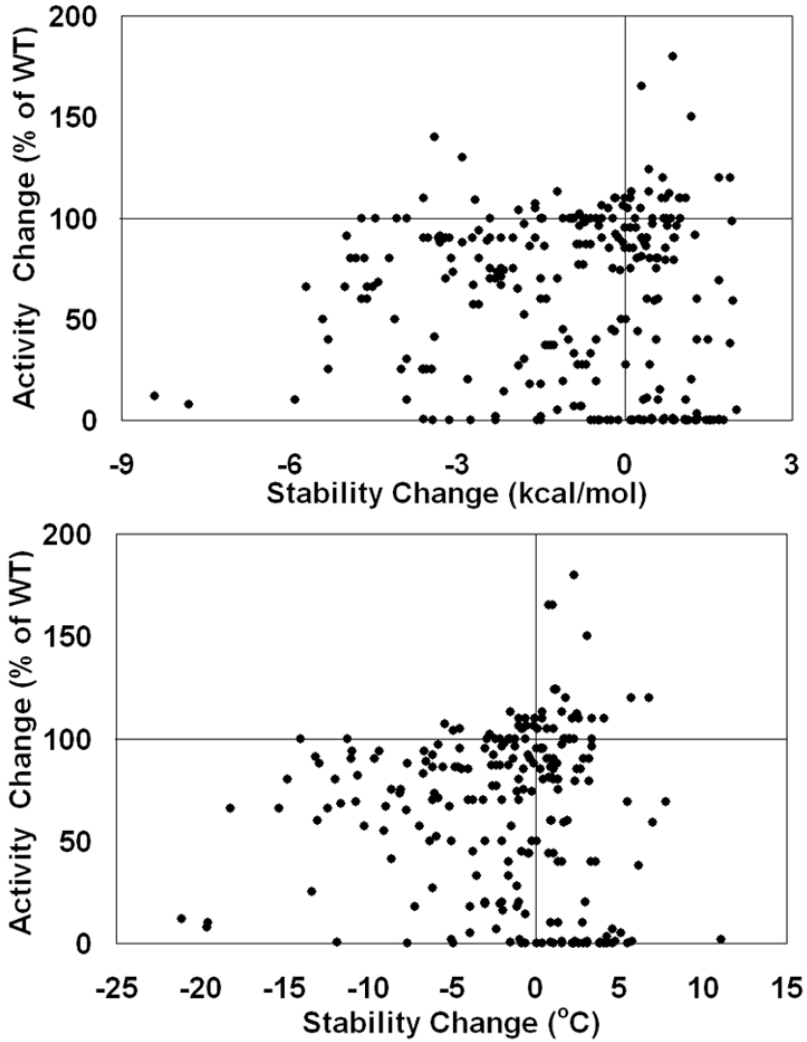
References:
D. Rennell *et al.*, *J Mol Biol* 222, 1991, pp. 67-88.
K. Saraboji *et al.*, *Comput Biol Chem* 29, 2005, pp. 25-35

Motivating Example: Phage T4 lysozyme

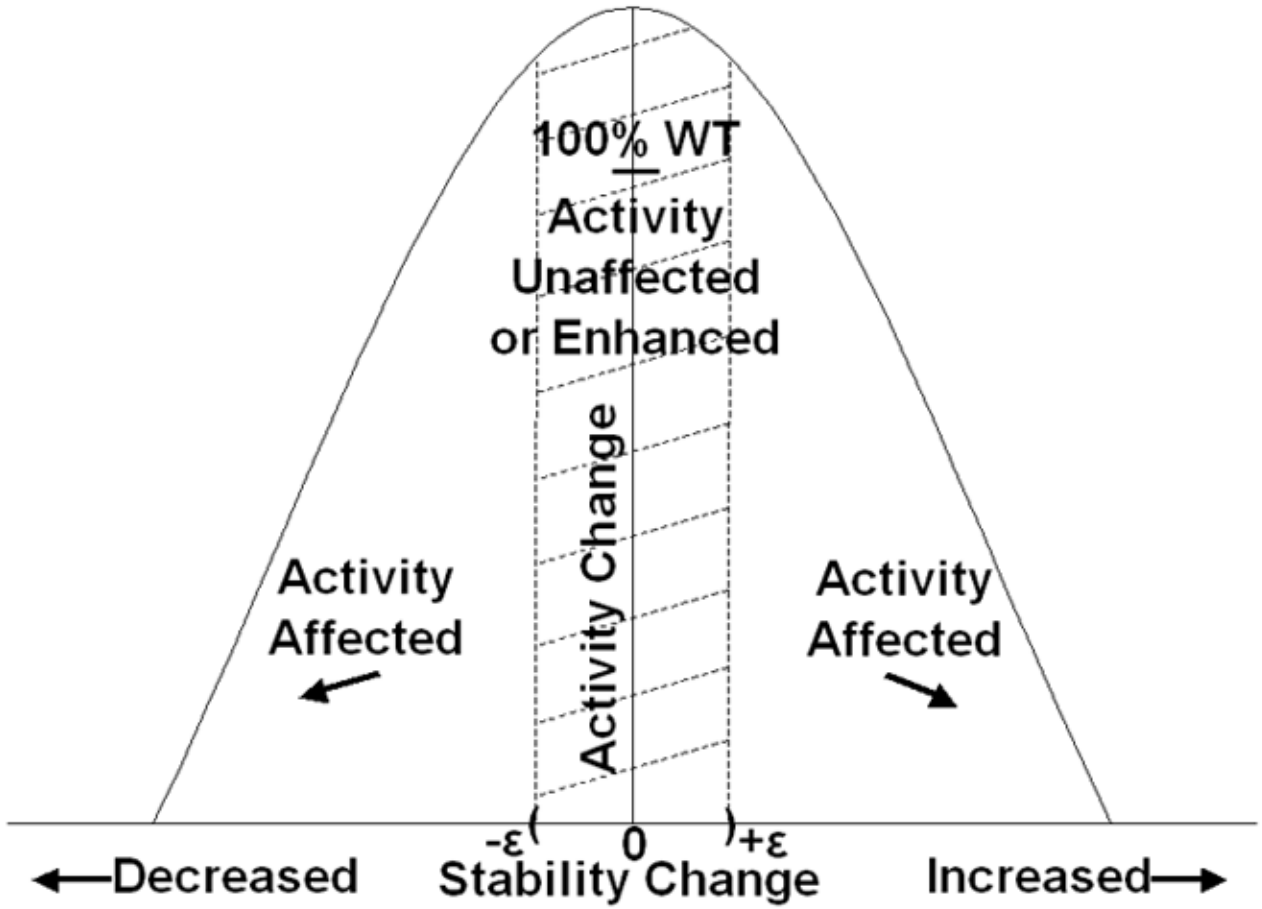


- Mutants in all four activity/stability change category pairs
- MRS > 0 for both inc categories, MRS < 0 for both dec categories
  - Reflects influence of sequence-structure compatibility on overall stability
  - MRS(inc) = 0.64, MRS(dec) = –0.87
  - Statistically significant structure (MRS) – stability (inc/dec categories) relationship (*t*-test, $p < 0.0005$)
- MRS(U) = –0.48, MRS(A) = –0.66 (trend not significant, too few mutants – 4%)

A Targeted ProTherm Search …



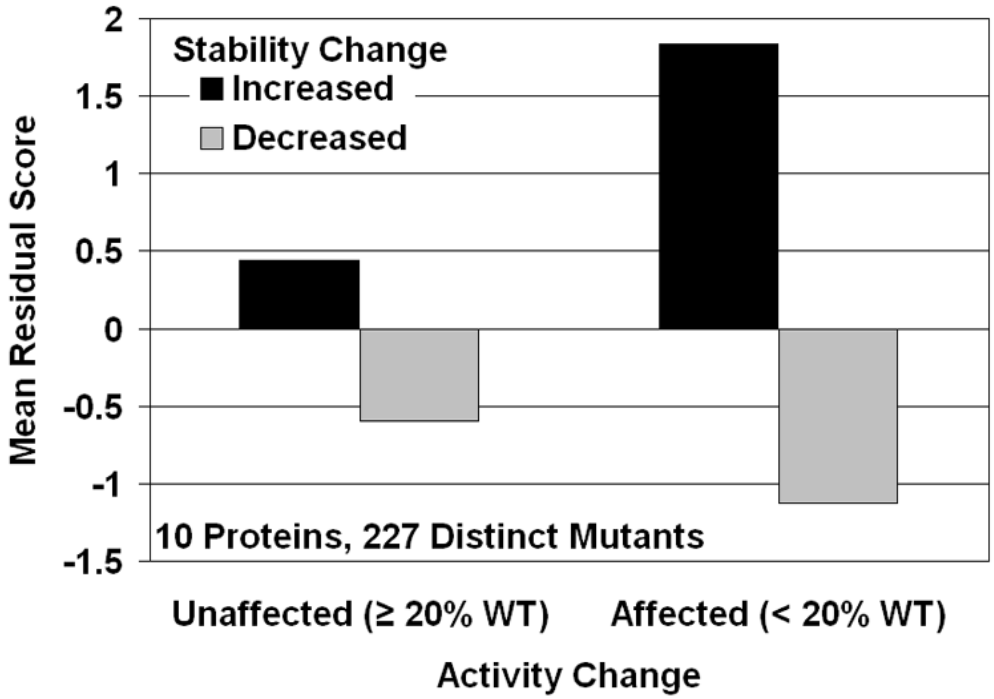
…Leads to a Stability-Activity Hypothesis



Experimental Data

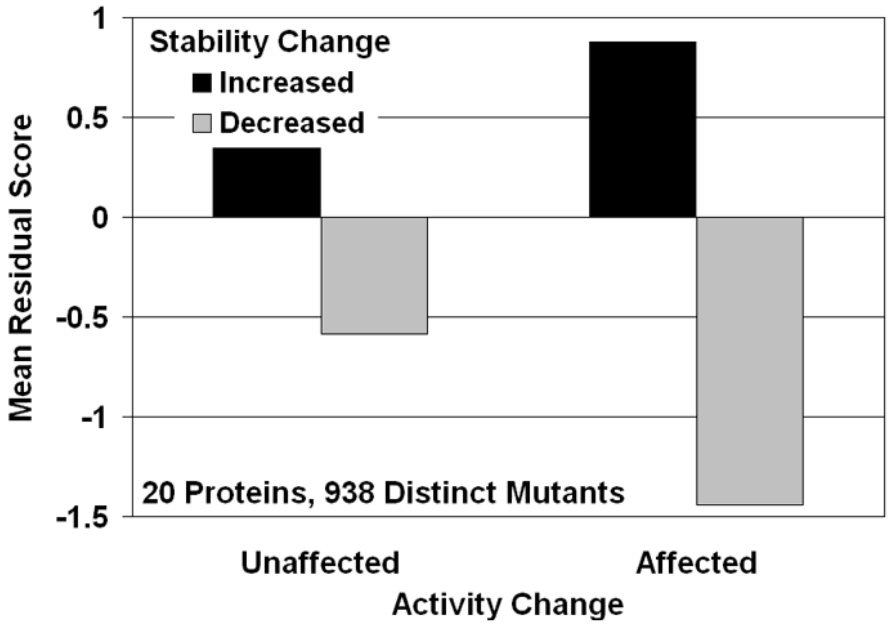
- ProTherm: 531 total mutants (includes repeats)
  - Both activity and stability change data available
  - 227 distinct mutants, from 10 proteins with known structures
- Our own literature search
  - 711 additional mutants, distinct from the ProTherm set
  - From 10 additional proteins with known structures
- Combined final dataset: 938 distinct mutants from 20 diverse proteins with known structures
- Complete dataset details, with mutant residual scores, available at: http://proteins.gmu.edu/automute/stability-activity.txt

Initial Result: ProTherm Mutants



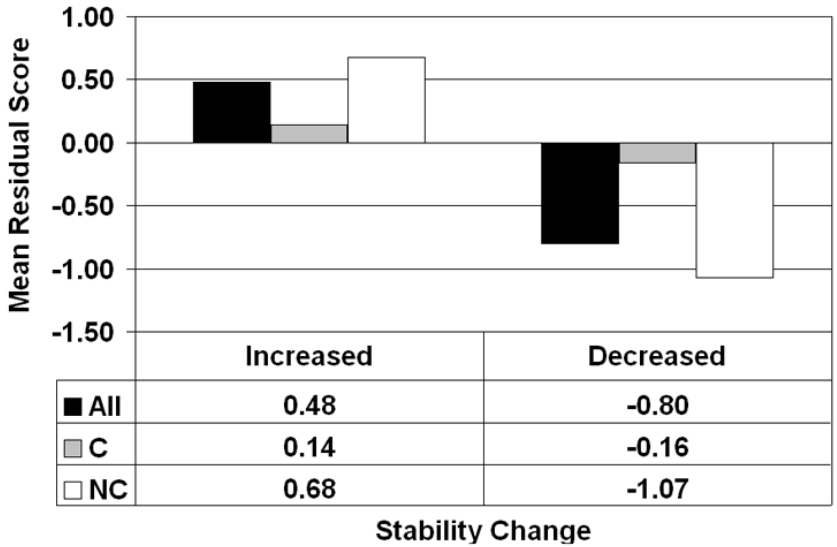
- MRS(inc) = 0.93, MRS(dec) = –0.68; statistically significant structure (MRS) – stability (inc/dec) relationship (*t*-test, $p < 0.0001$)
- structure (MRS) – function (U/A) relationship is not evident here …next step, investigate whether small sample size is the reason.

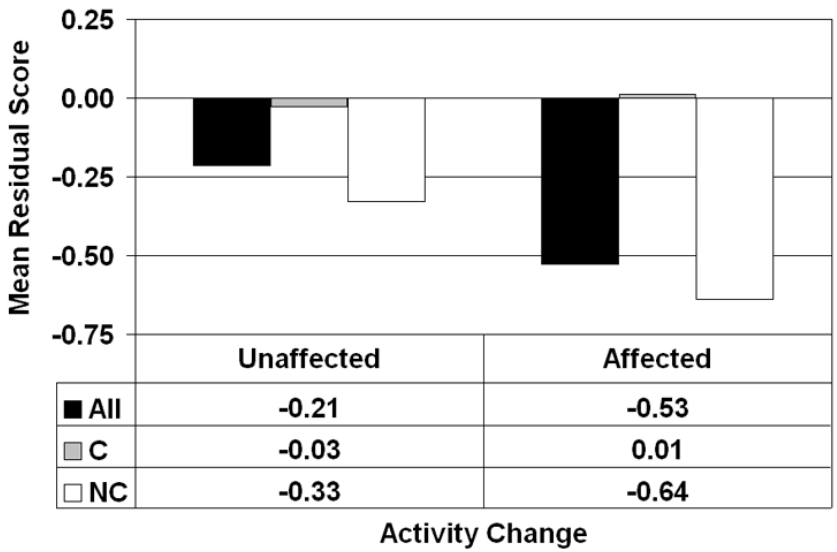
Final Result: Combined Dataset of 938 Mutants



- Subset sizes: 279 U\inc, 421 U\dec, 94 A\inc, and 144 A\dec
- statistically significant structure (MRS) – stability (inc/dec) relationship (*t*-test, $p < 0.0001$) and
- statistically significant structure (MRS) – function (U/A) relationship (*t*-test, $p < 0.05$)

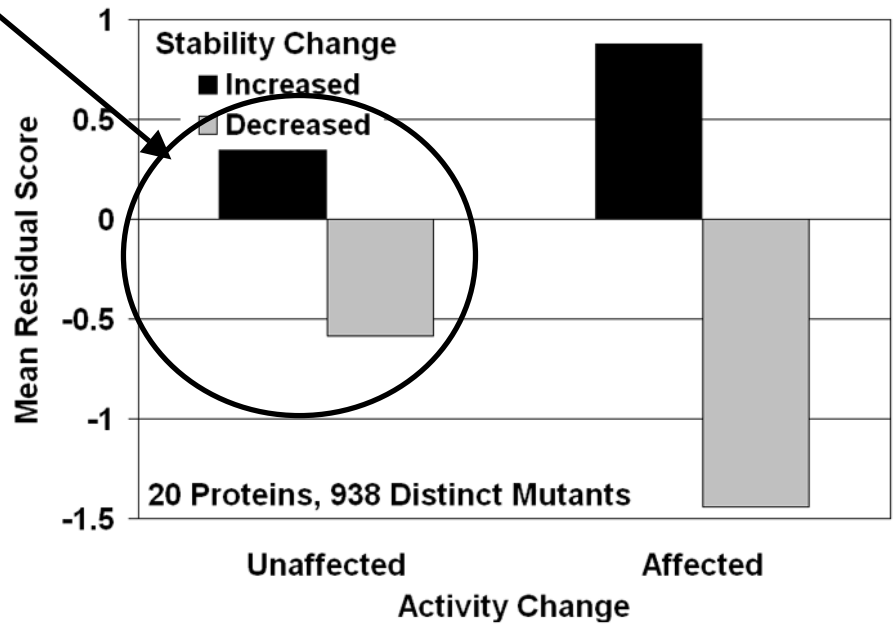
Combined Dataset: Details





Conclusion

- Distinctive trend evident in the stability-activity-structure plots is a consequence of the following general principles…
- First, mutations having minimal impact on protein sequence-structure compatibility (MRS of small magnitude) also have minimal effect on stability (inc/dec) and leave activity unaffected



Conclusion

- Next, two reasons why activity would be detrimentally affected:
  1. Mutant sequence-structure compatibility (MRS) significantly increases relative to the native protein, corresponding to mutants that are highly stable (i.e., too rigid to accommodate substrates or catalyze reactions)
  2. MRS significantly decreases relative to the native protein, corresponding to mutants that are highly unstable (too flexible due to lack of a sufficient noncovalent bonding network to maintain the proper fold)