

Model Testing for a Time Dependent Transition of a Rates Across Sites to a Covarion Model of Protein Evolution

Makayla Tisdell, David Liberles
Department of Molecular Biology
University of Wyoming



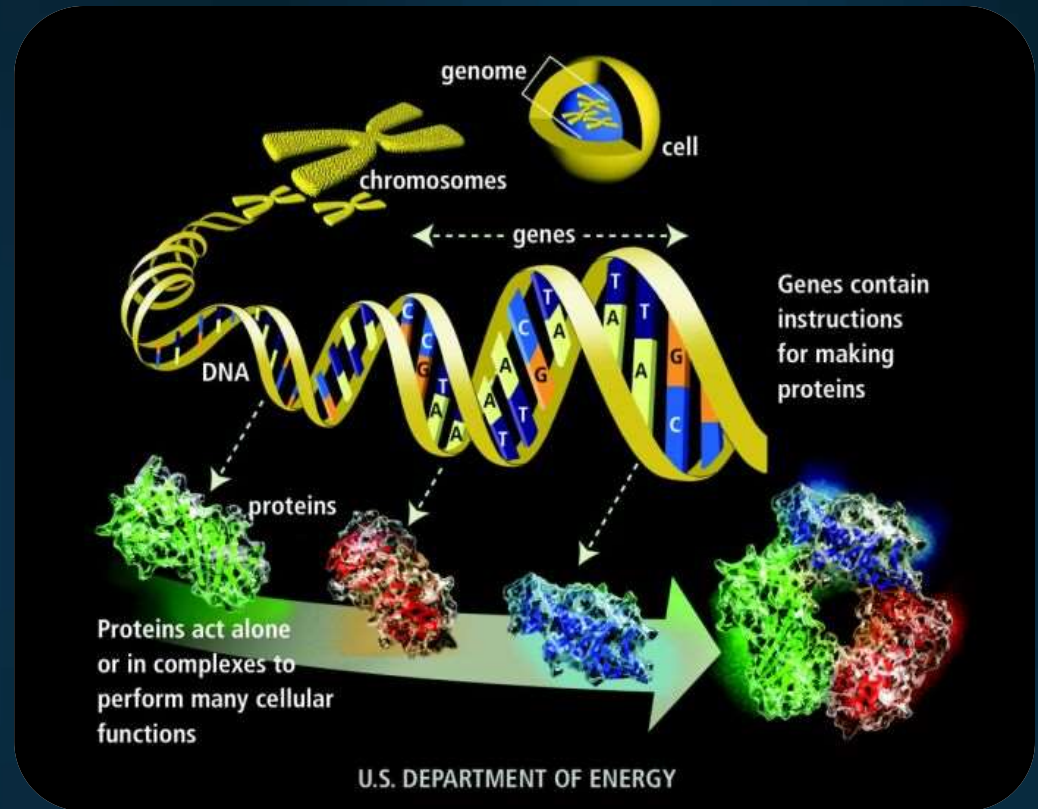


Abstract

Protein-coding gene sequences typically evolve constrained by the requirements for a protein to fold into its three dimensional structure. These constraints can dictate evolutionary rates at different sites, where residues in the hydrophobic core of a protein typically evolve more slowly than those on the surface. The exceptions to this are surface residues involved directly in functions of the protein such as binding, which are also conserved. A gamma distribution of rates across sites is typically used to describe the process of protein evolution in what is called the rates across sites model. It has been proposed that this model is violated when the function of a protein changes. It has alternatively been proposed that protein structure leads to violations of the gamma distribution over increasing evolutionary time. I hypothesized that the rate of the transition is fold-dependent and that different protein folds will move from an RAS model to a Covarion model at different rates. In preliminary analysis, an equal rates rather than a RAS model was supported which suggests the energy function used in the evaluation of protein folding should be modified before further analysis.

Molecular Evolution

Evolution at the gene sequence level



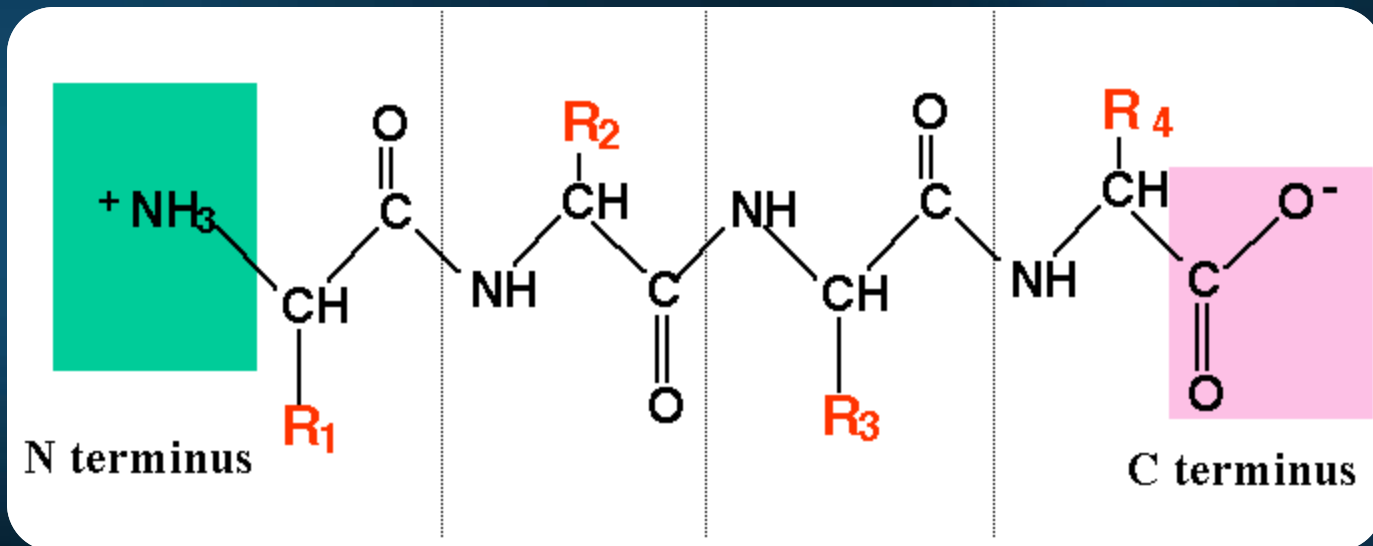
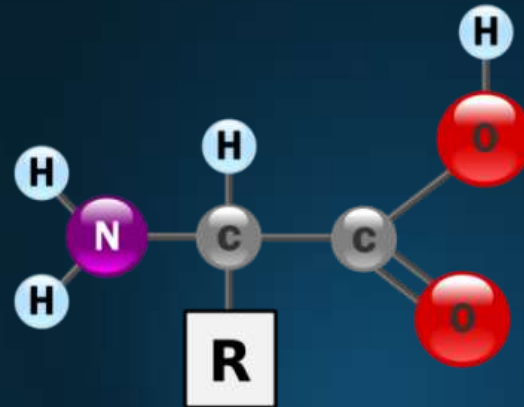
U.S. DEPARTMENT OF ENERGY

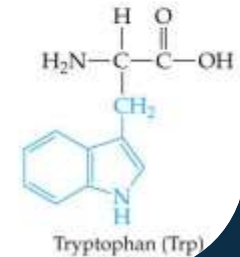
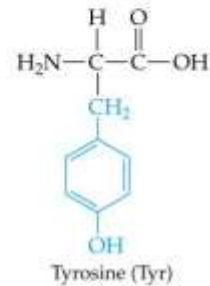
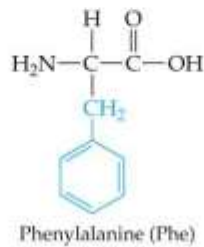
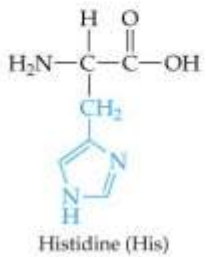
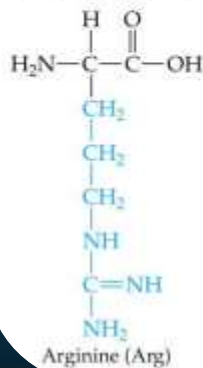
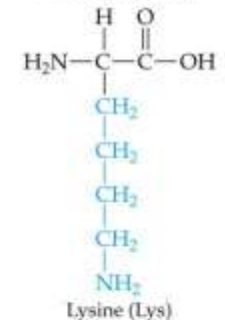
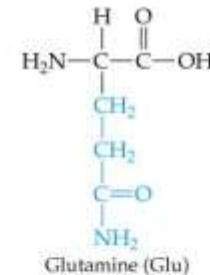
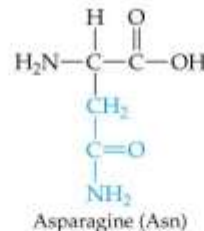
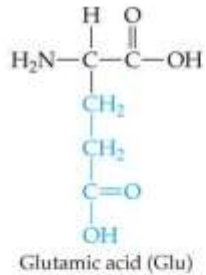
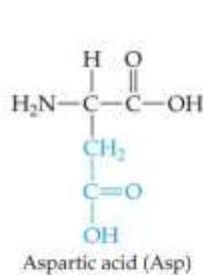
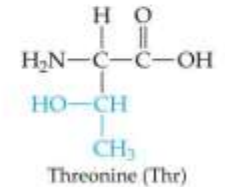
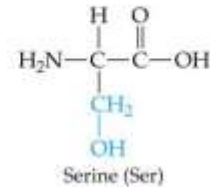
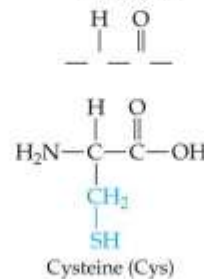
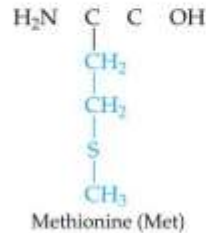
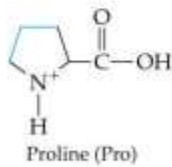
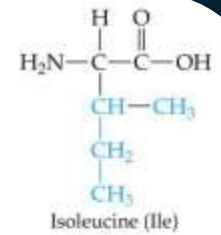
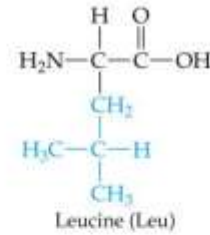
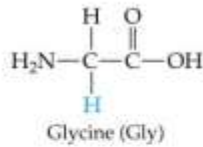
U.S. DEPARTMENT OF ENERGY

functions
perform many cellular
or in complexes to
perform many

Amino Acids

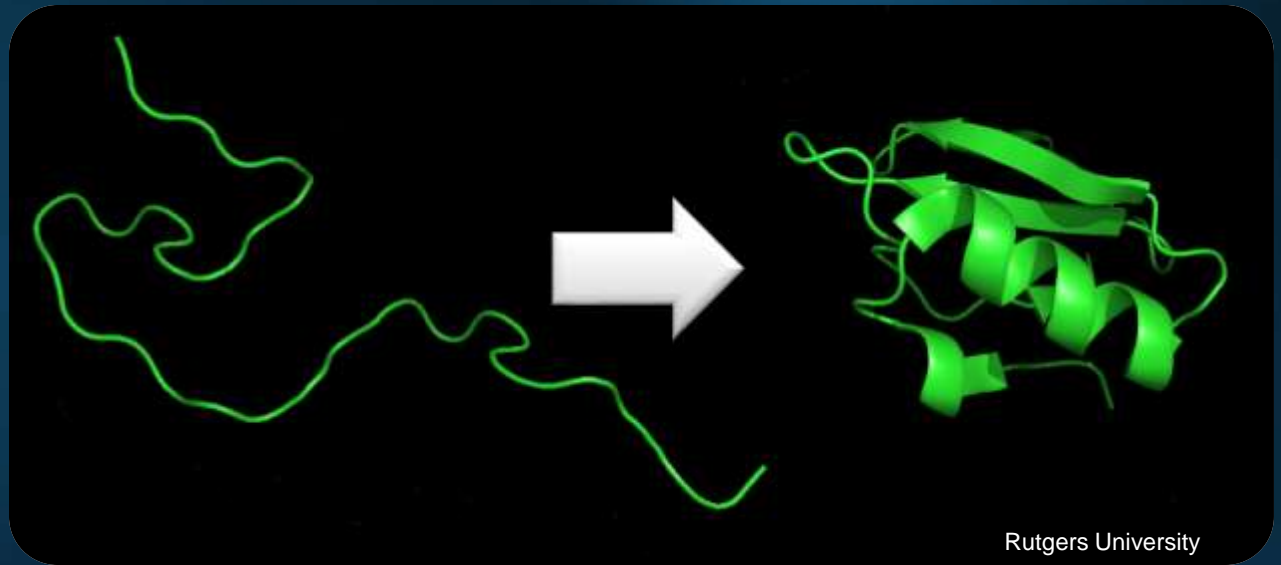
- General structure
- “R” Group denotes side chain
- The physicochemical properties of the side chain place constraints on protein evolution





Protein Folding

$$\Delta G = \Delta H - T\Delta S$$



Mutation effect on protein folding

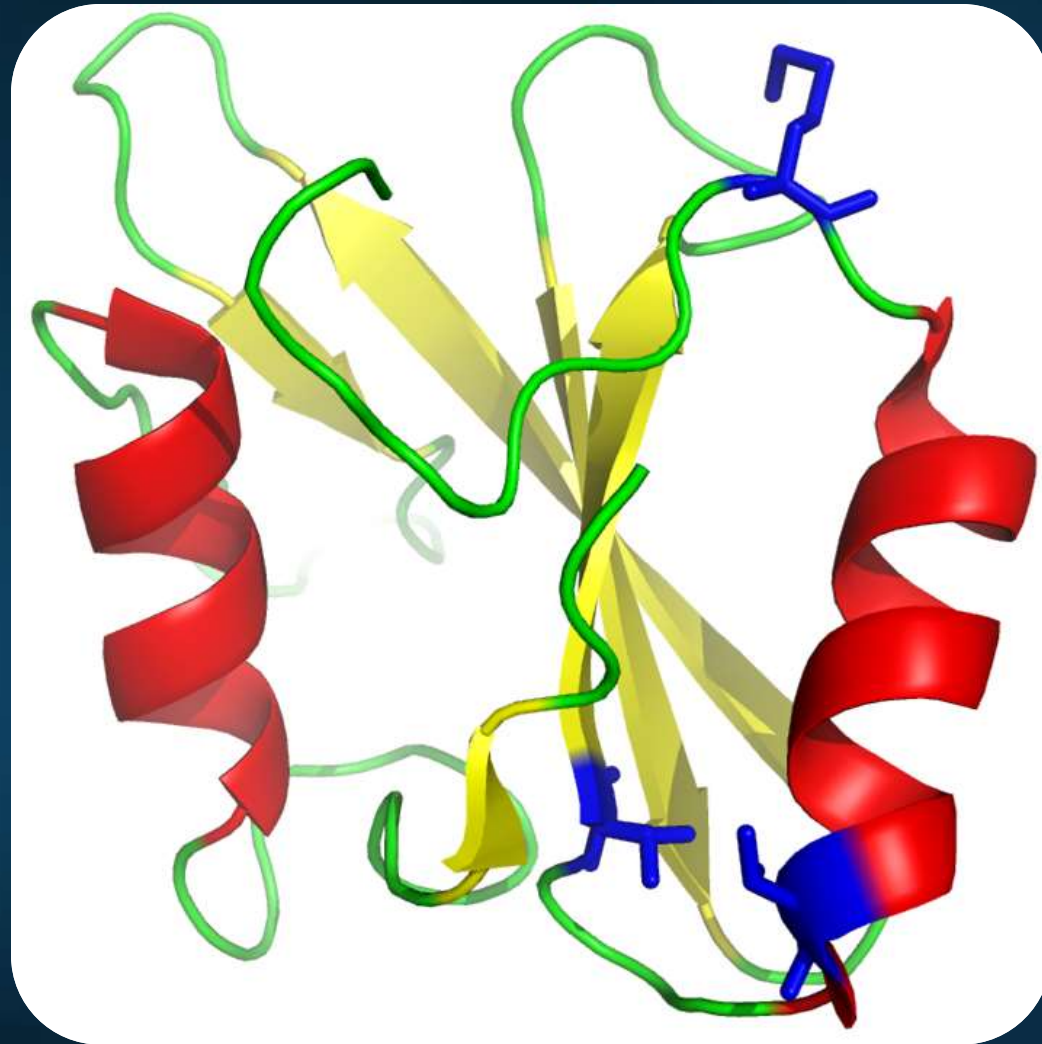


Image
courtesy
Johan
Grahnen

Introduce a mutation

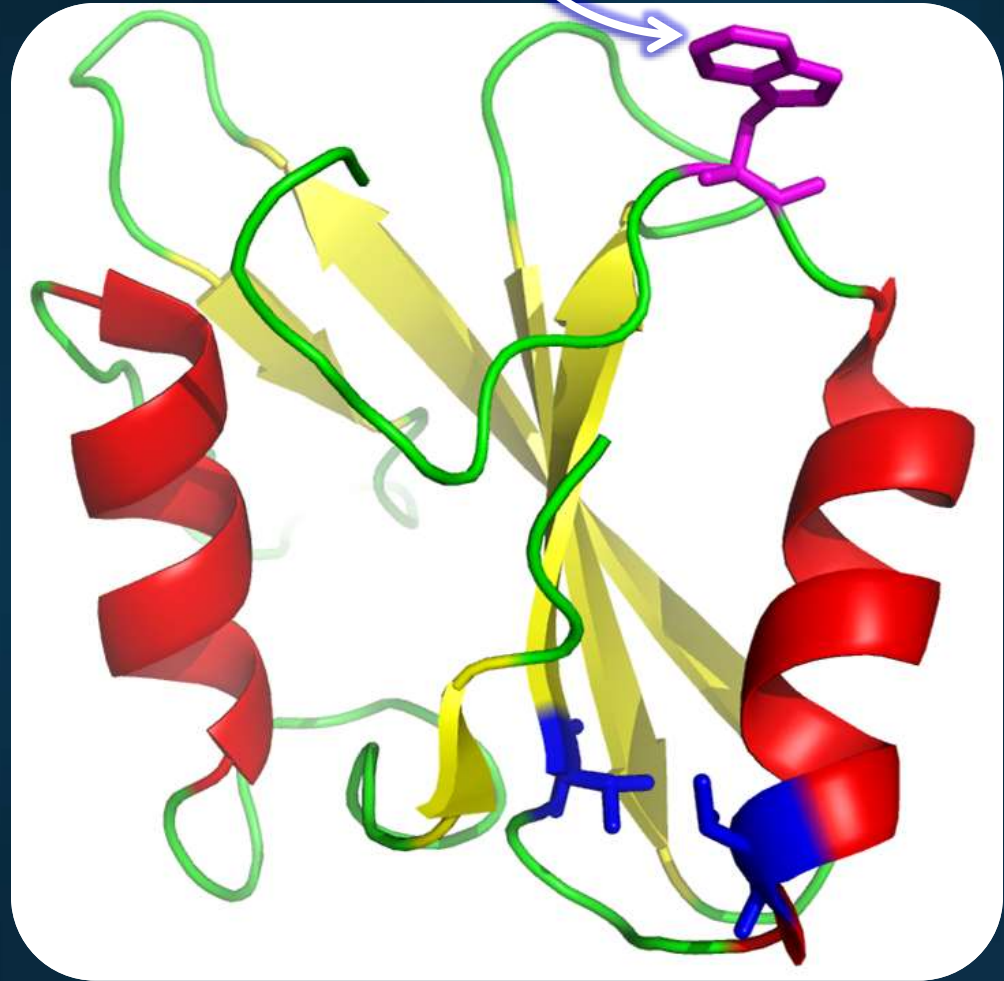


Image
courtesy
Johan
Grahnen

An additional mutation

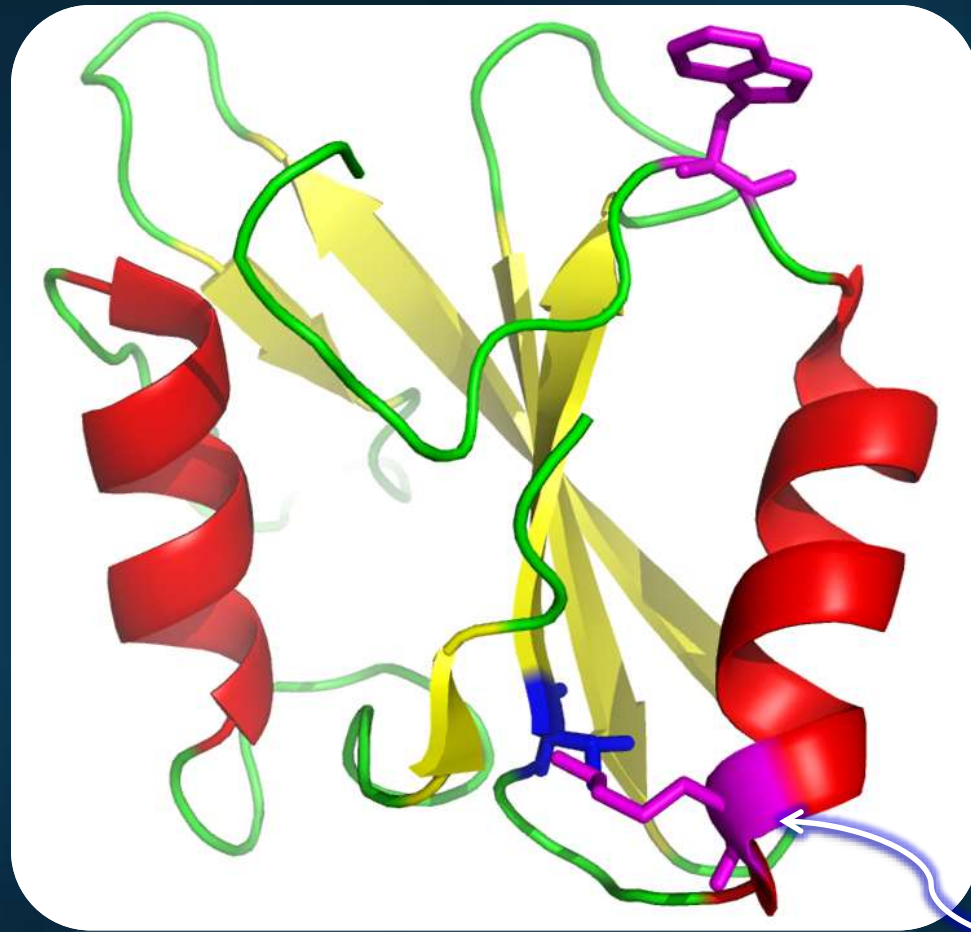


Image
courtesy
Johan
Grahnen

Positive Selection

- Non-synonymous substitutions are important in evolution.
- New functions in a protein can be potentially created by substitution of functionally significant amino acids
- Understanding the evolution of functional characteristics among members of the same protein family is important.
- Positive selection occurs when an advantageous non-synonymous substitution sweeps a population.

Mutational Change

Structural Change

Functional Change



Modeling Protein Evolution

- Challenging
- Increasing degree of complexity in models
 - Equal Rates
 - Rates across Sites
 - Rate shifting (covarion)
- Prediction of functional change and the debate that followed
 - Gu
 - Philippe

Research Hypotheses

- Hypothesis I
 - There is a time-dependent transition between the rates across sites model and the covarion models that occurs without a functional shift.
- Hypothesis II
 - The rate of this transition depends on specific protein folds





Research Plan



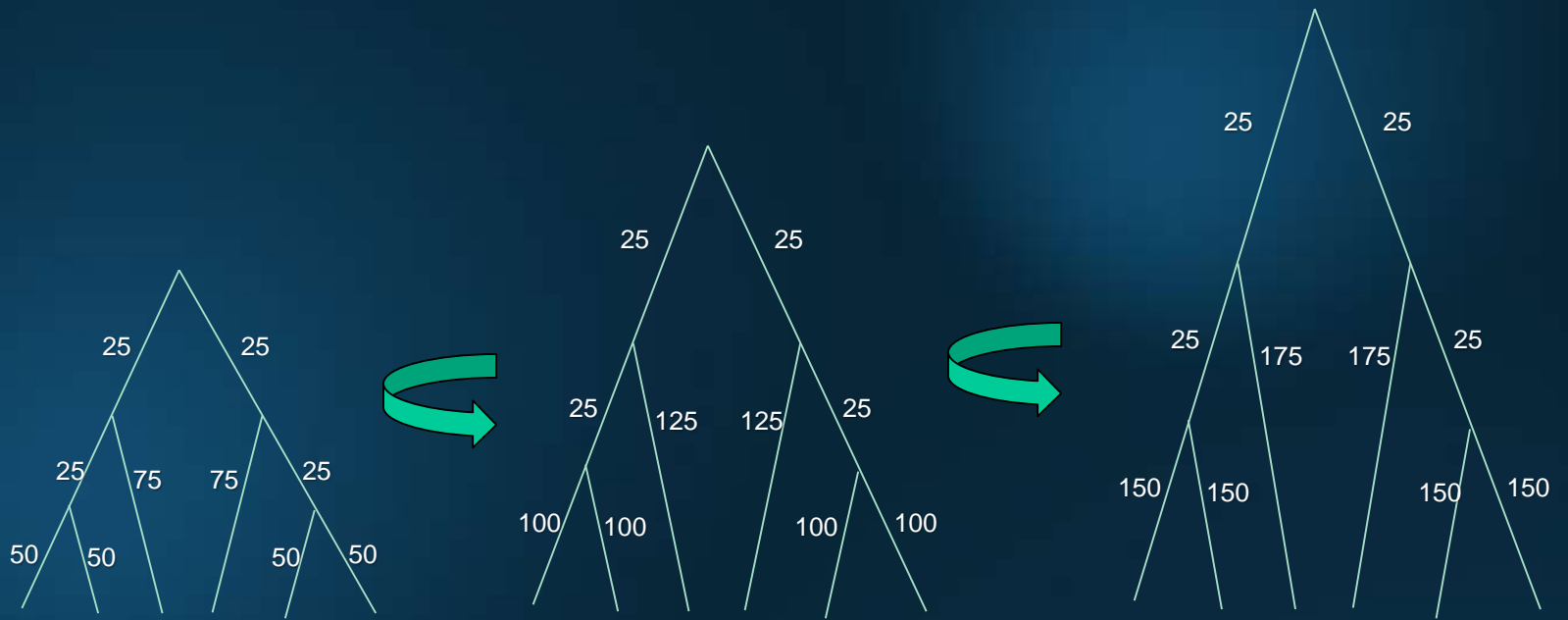
Tree
Topology
and Protein
Sequence

Simulation
with
Johan's
Program

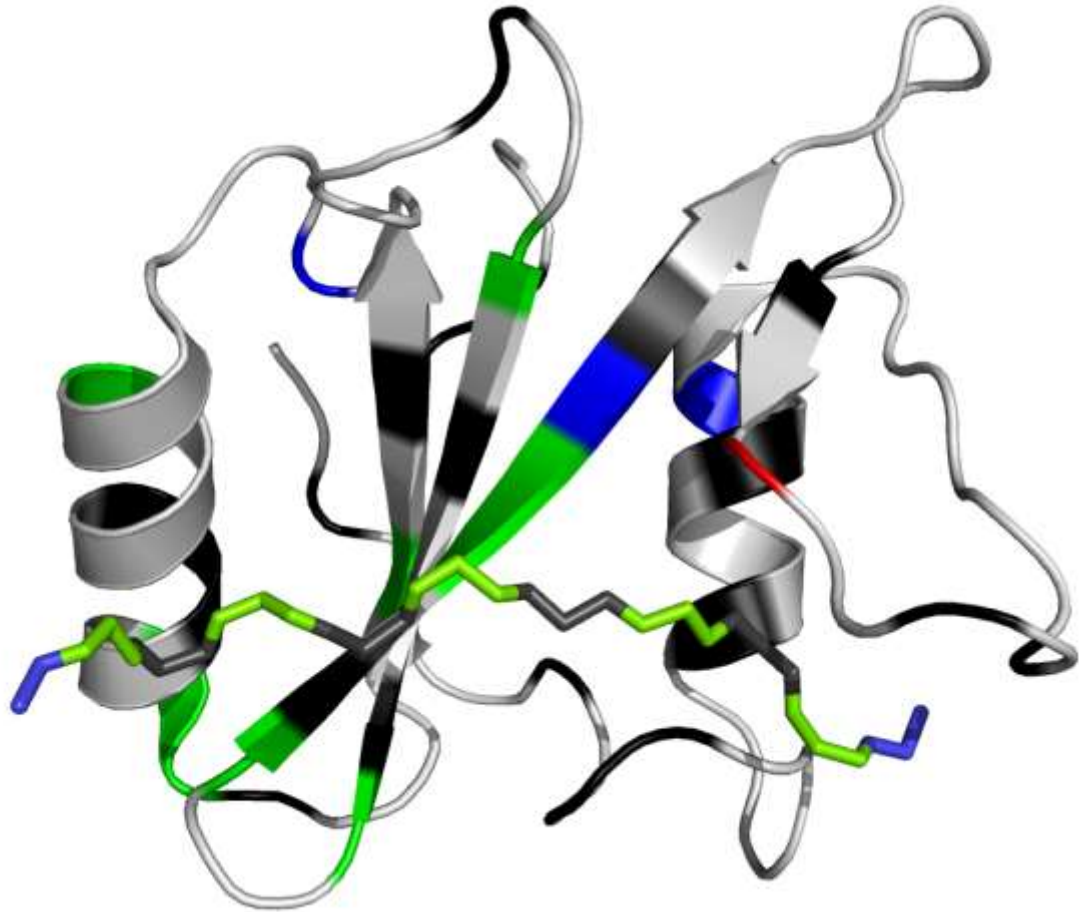
ProtTest
and Procov

Entropy
and
Divergence

Tree Topology



1D4T (SH2 domain)





Simulation strategy

Sequence known to fold and bind to ligand

- SH2 domains

Mutate the sequence

Evaluate the effect of the mutation

- Folding, stability
- Binding affinity and specificity

Characterize the viability



ProCov and ProtTest

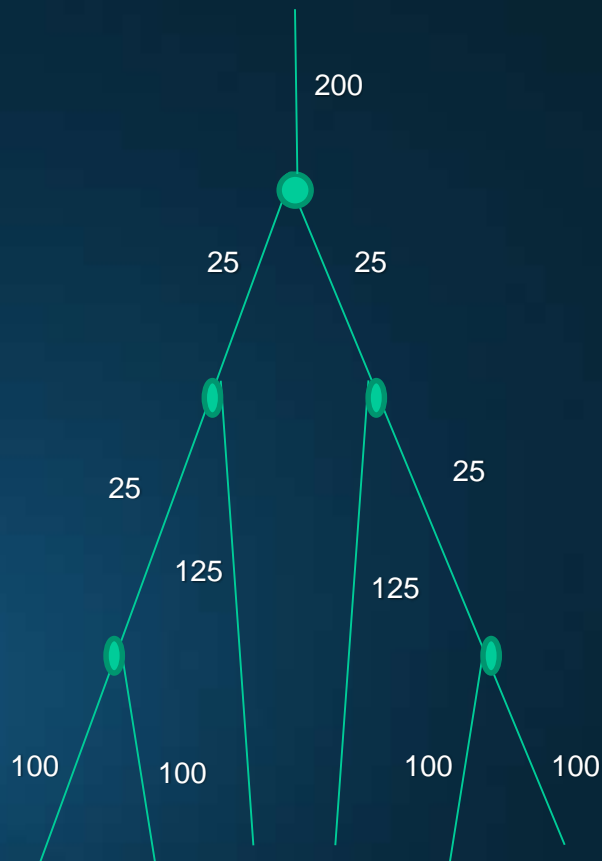
- Alignment
- Tree
- Procov
 - Covarion models (Galtier, Tuffley, Huelsenbeck, General) vs. a RAS model.
- ProtTest
 - Selects the model of protein evolution that best fits a given set of sequences.



Results from negative selection

- Tree1: HIVb + F
- Tree2: HIVw + F
- Tree3: JTT + F
- Created equal rates model to test with procv
- By use of AIC criterion, the equal rates model best fit the data

Not what I was looking for...



Simulated with negative selection over reformatted tree.

Ran 5 replications of the simulation

Ran a 200 generation to equilibrate before branching and increased terminal branches by 50 generations.



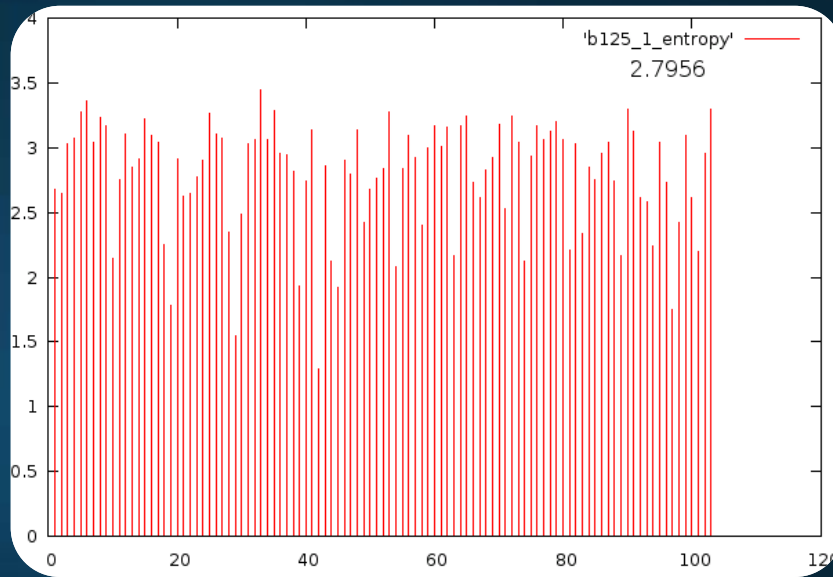
Results?

- Rep1
 - tuffley > huelsenbeck > RAS > equal rates > general
- Rep 2
 - equal rates > tuffley > RAS > huelsenbeck > general
- Rep 3
 - equal rates > RAS > tuffley > huelsenbeck > general
- Rep 4
 - equal rates > RAS > tuffley > huelsenbeck > general
- Rep 5
 - equal rates > RAS > tuffley > huelsenbeck > general
- Still no support for gamma in ProtTest

Analysis of Results

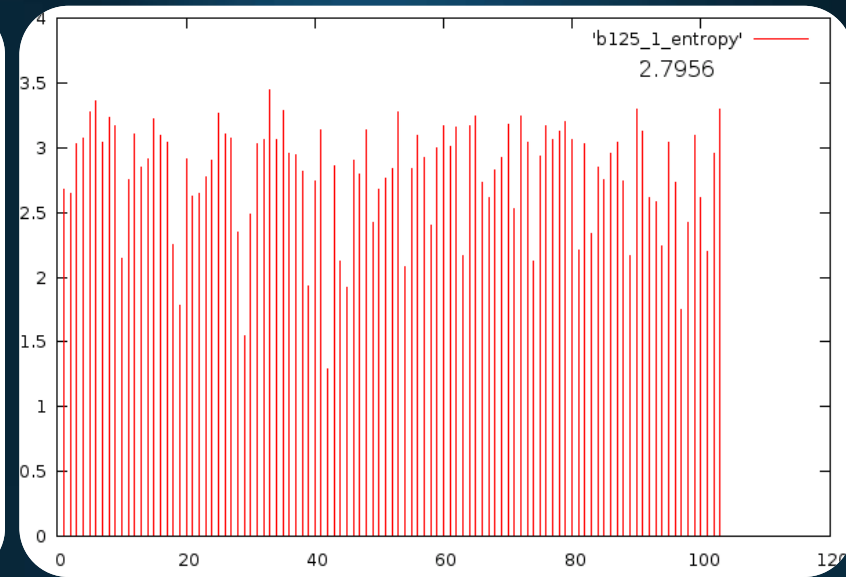
Entropy

6 terminal branches



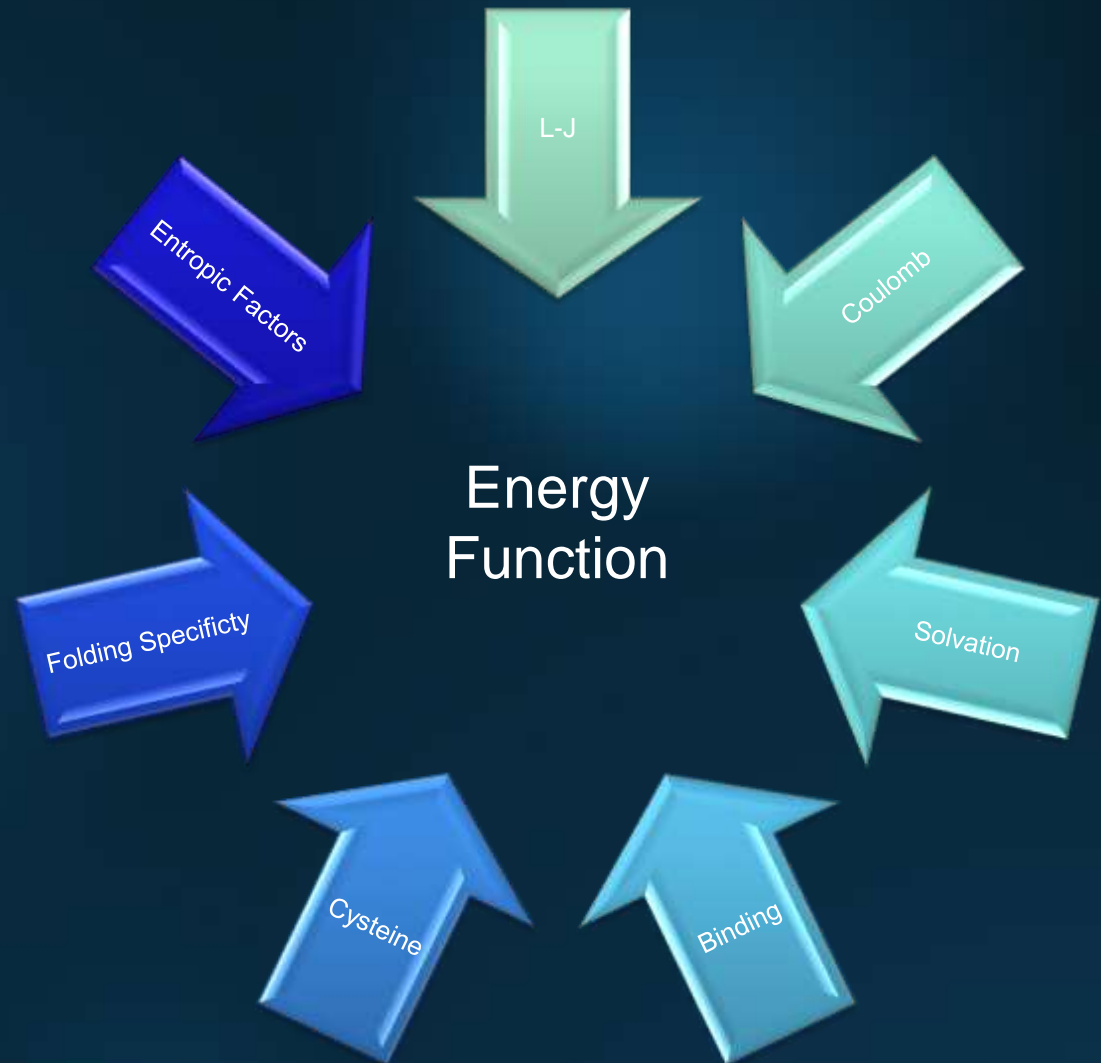
Compare conserved sites

42



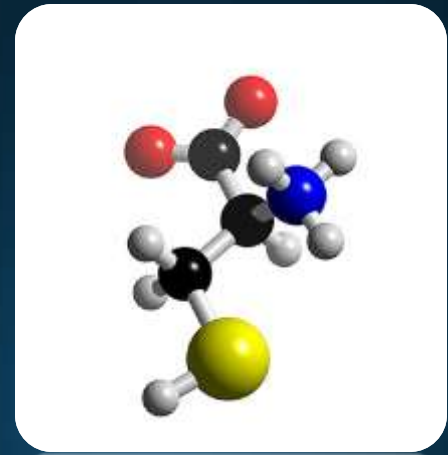
Conclusions and Forward Steps

- Energy Function
- Proper model
- Decoy Ligands
- Cysteine Term

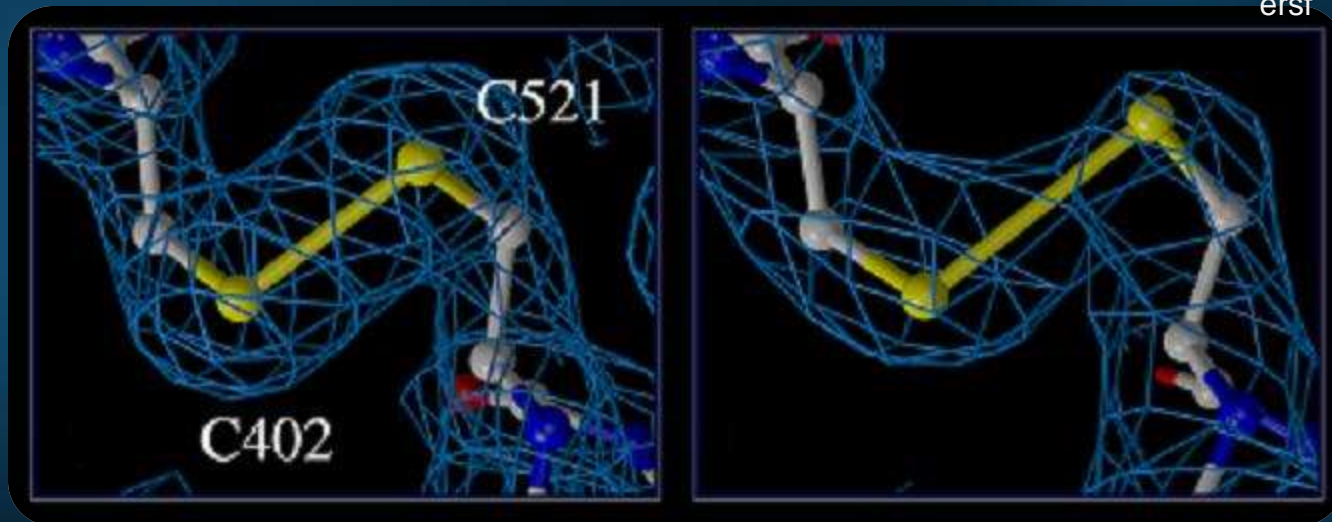


Disulphide bridges

- Distance
- Angle
- Redox potential



ersf





Resources

- Philippe H, Casane D, Gribaldo S, Lopez P, Meunier J. **Heterotachy and functional shift in protein evolution.** *IUBMB Life*. 2003 Apr-May;55(4-5):257-65.
- Gu, Xun, Eric A. Gaucher, Michael M. Miyamoto, and Steven A. Benner. "Predicting Functional Divergence in Protein Evolution by Site-specific Rate Shifts." *Trends in Biochemical Sciences* 27.6 (2002).
- Wang H-C, Susko E. & A. J. Roger, *PROCOV*
- PROTTEST: Selection of best-fit models of protein evolution sponsored by © 2004-2009 Federico Abascal, Rafael Zardoya and David Posada 2.4 (October 09).
- weblogo.berkeley.edu/logo.cgi

- Special thanks to Johan Grahnen and David Liberles for all their help
- Thanks to INBRE for funding!

Questions?

