

KINARI-Mutagen Case Study: Validation Using 48 Mutants

Filip Jagodzinski

25 April 2012

To determine if KINARI-Mutagen could correctly identify destabilizing mutations in a wider range of proteins, we searched the ProTherm Database[1], which catalogues $\Delta\Delta G$ values for substitutions that have been performed in the physical protein. A total of 167 entries had mutations to glycine. Of those, 48 mutants among 14 proteins had single-point substitutions. We also chose PDB files that had all core residues resolved.

We used KINARI-Mutagen to generate the 48 *in silico* mutants and analyze their rigidity. Along with the SASA value for each wild-type residue at the location where the mutation was performed, we tallied the change to the largest rigid body of the protein caused by the point mutation, and the degree of hydrophobicity of each wild-type residue, using the Kyte and Doolite hydrophobicity scale[2]. The output of KINARI was also used to tally how many hydrogen bonds and hydrophobic interactions were lost due to the mutation. To facilitate analysis, the 48 mutants were grouped according to whether the substituted residue engaged in hydrogen bonds and hydrophobic interactions (Table 1).

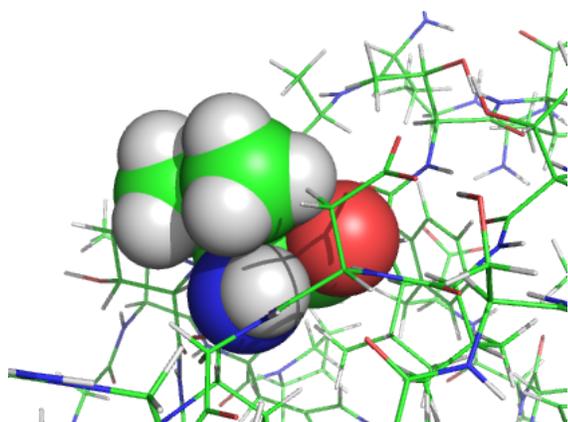
KINARI-Mutagen relies on the loss of hydrogen bonds and hydrophobic interactions upon a residue's change to glycine, to predict the effects of a mutation. Thus we did not expect to accurately predict a substitution as destabilizing, if KINARI found that in the wild-type protein the amino acid engaged in neither hydrogen bonds nor hydrophobic interactions (Group 1). Group 2 has entries for which the residue of the wild-type protein was solvent exposed (more than 50% of the residue was exposed). Because these residues are on the periphery of the protein, their being mutated to glycine would not be expected to have a large effect on the size of the largest rigid cluster, especially if the side chain of the residue was protruding fully into the solvent (Figure 1(a)).

In Group 3, four mutants had wild-type amino acids (Valine, Leucine, Methionine, Phenylalanine) that do form hydrophobic interactions that can be observed by visual inspection. However, because of the packing of these core residues in this structure which were slightly less tight than in many protein cores, the algorithm in KINARI to detect hydrophobic interactions detected far too few of them. Figure 1(b) shows a phenylalanine that upon visual inspected should have been stabilized via several hydrophobic interactions, but no atoms in the residue were within 3.5Å of a

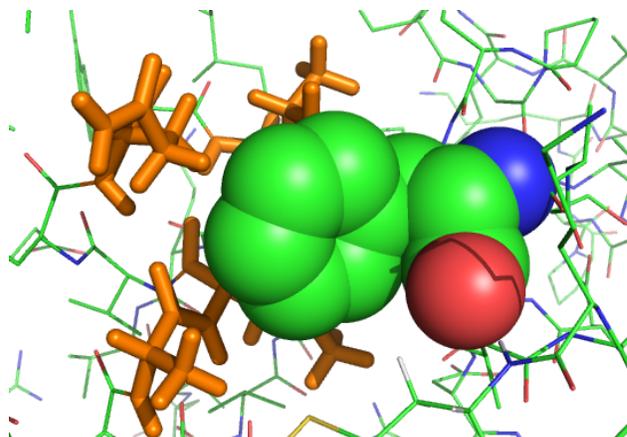
Group	Description of Wild Type AA at Mutation Point	# Mutants	Identified As Destabilizing
1	No hydrogen bonds or hydrophobics detected	13	0
2	Solvent exposed (>50%)	8	0
3	Too few hydrophobic	4	0
4	Stabilizing interactions found	23	22

Table 1: The rigidity results for the 48 mutants was grouped based on whether the mutated residue engaged in stabilizing interactions.

heavy neighbor atom, so no hydrophobic interactions were detected. Had the atoms of that part of the structure been oriented slightly differently to allow closer packing, KINARI's hydrophobic detection algorithm would have placed more hydrophobic interactions there, which could have caused that residue to be labeled as critical when it was mutated.



(a) In the *Streptomyces Subtilisin Protease inhibitor*, (PDB ID 3ssi), Valine 13 (spheres) is 56% exposed, so only 1 hydrophobic was detected, precluding KINARI analysis.



(b) In *Staphylococcal Nuclease*, (PDB 1stn) Penylalalanine 61 (green spheres) lies in a hydrophobic pocket (orange), but no hydrophobics were detected.

Figure 1: Some amino acids that are highly solvent exposed were not identified as destabilizing, because they did not engage in stabilizing interactions (Figure 1(a)). Some residues like those that are completely or largely solvent inaccessible lie more than 3.5\AA from the nearest heavy atom, so hydrophobic interactions (orange sticks, Figure 1(b)) in the range of 3.6\AA to 4.5\AA are not found by the hydrophobic detection algorithm in KINARI, preventing quantitative analysis of the impact of the mutation to glycine.

Group 4 contains 23 mutants that had more reasonable numbers of hydrophobic interactions which were identified by KINARI, and many of them have hydrogen bonds. Of these, 22 were identified as critical, based on the fact that these mutants had largest rigid clusters that were smaller than the largest rigid cluster of the wild-type protein.

From the analysis of these 48 mutants, this first implementation of KINARI-Mutagen is able to make qualitative stability predictions. In the cases when residues are highly solvent exposed, KINARI-Mutagen is not as accurate, because such residues do not engage in as many stabilizing interactions as would be expected of them. Similarly, the pre-existing algorithm to detect hydrophobic interactions is not always accurate, when compared to the predicted hydrophobic interactions from a visual inspection. In future work, we plan to address this hydrophobic interaction algorithm.

References

- [1] M.D. Kumar, K.A. Bava, M.M. Gromiha, P. Prabakaran, K. Kitajima, H. Uedaira, and A. Sarai. Protherm and pronit: thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic Acids Research*, 34(suppl 1):D204–D206, 2005.
- [2] J. Kyte and R.F. Doolittle. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1):105–132, 1982.