Title: AN ATOMIC VIEW OF ADDITIVE MUTATIONAL EFFECTS IN A PROTEIN STRUCTURE

Author(s): M.M. Skinner, T. C. Terwilliger

Submission to: Hanford Symp./Health & the Environment, Hanford, WA
October 24-26, 1995

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
An Atomic View of Additive Mutational Effects in a Protein Structure

Matthew M. Skinner
and
Thomas C. Terwilliger

Structural Biology Group
Mail Stop M888
Los Alamos National Laboratory
Los Alamos, NM 87545

Tel: (505) 667-0072
Fax: (505) 665-3024
email: terwilliger@LANL.gov

Running head: Additivity of mutational effects

Keywords: Protein crystallography/protein structure/protein engineering/protein-nucleic acid interactions
Abstract

Substitution of a single amino acid in a protein will often lead to substantial changes in properties. If these properties could be altered in a rational way then proteins could be readily generated with functions tailored to specific uses. When amino acid substitutions are made at well-separated locations in a single protein, their effects are generally additive. Additivity of effects of amino acid substitutions is very useful because the properties of proteins with any combination of substitutions can be inferred directly from those of the proteins with single changes. It would therefore be of considerable interest to have a means of knowing whether substitutions at a particular pair of sites in a protein are likely to lead to additive effects.

The structural basis for additivity of effects of mutations on protein function was examined by determining crystal structures of single and double mutants in the hydrophobic core of gene V protein. Structural effects of mutations were found to be cumulative when two mutations were made in a single protein. Additivity occurs in this case because the regions structurally affected by mutations at the two sites do not overlap even though the sites are separated by only 9 Å. Structural distortions induced by mutations in gene V protein decrease rapidly, but not isotropically, with distance from the site of mutation. It is anticipated that cases where structural and functional effects of mutations will be additive could be identified simply by examining whether the regions structurally affected by each component mutation overlap.
Additivity of mutational effects on protein function

It is now widely appreciated that effects of mutations on a property of a protein molecule are sometimes additive (Wells, 1990). This means, for example, that if a particular mutation increases the stability of the protein by 1 kcal/mol, and if a second mutation increases the stability by 2 kcal/mol, then a protein containing both mutations might be stabilized by about 3 kcal/mol. In cases studied so far, additivity of this type is generally observed if the amino acid side chains that are altered in the two mutants are not in direct contact with each other and the amino acid side chains involved are uncharged (Hecht et al., 1986; Matsumura et al., 1989; Sandberg and Terwilliger, 1991; Stearman et al., 1988; Wells, 1990; Zhang et al., 1991). Our laboratory provided a demonstration of this with a set of 16 double mutants at two sites in the gene V protein in which the two sites were not in direct contact (Sandberg and Terwilliger, 1991). Other examples where the effects of mutations are nearly additive include effects on DNA binding specificity of λ repressor (Nelson and Sauer, 1985), substrate specificity of subtilisin (Carter and Wells, 1987; Wells et al., 1987a; Wells et al., 1987b; Wells et al., 1987c), and cofactor binding affinity of glutathione reductase (Scrutton et al., 1990).

Recently the deviations of effects of multiple mutations at well-separated sites from exact additivity in a number of proteins were examined in detail (LiCata and Ackers, 1995). In several cases a pattern of either super- or sub-additivity was found, suggesting some systematic kind of interaction among the mutated residues, although the nature of this interaction is unknown. It appears in any event that small non-additive effects might result from mutations at distant sites. A special case of such non-additive effects were the mutants of staphylococcal nuclease in which non-additive effects on protein stability were evidently due to interactions occurring in a partially-folded or unfolded state of the protein (Green and Shortle, 1993). In contrast to well-separated mutations, mutations at sites that
are in contact with one another often show strongly non-additive behavior (Hurley, 1994; Hurley et al., 1992; Mildvan et al., 1992). For example, altering the side chains facing each other at positions 33 and 47 in gene V protein, increasing the bulk of one and decreasing the bulk of the other, results in strongly sub-additive effects on stability (Sandberg and Terwilliger, 1993).

Using additivity to simplify the engineering problem

Single mutations in a protein generally affect many different properties of that protein. This often complicates engineering of a protein, where changing just one property may be desirable, or where secondary effects of a mutation (such as decreased stability, for example) render the desirable effects (such as increased affinity for a substrate) less useful. If mutations affect several different properties of a protein in an additive fashion, however, then combinations of mutations could be constructed by rational design that would affect one property of a protein without substantially affecting another (Sandberg and Terwilliger, 1993). This could be used to engineer the specificity of gene V protein-DNA interactions, for example, while simultaneously maintaining the stability of the protein. In effect, combinations could potentially be found in which the individual mutations cancel each other's undesirable effects, while adding to each other's desired effects. If effects on several different properties of a protein are all additive, then the properties of a mutant with several mutations can be thought of as a "vector sum" of the properties of the wild-type, where the vectors represent the changes in these properties from each mutation. This observation is quite important because it means that if we had a large "basis set" of mutants that were characterized with respect to several properties, we could design a new mutant with a desired combination of these properties, just by choosing a set of mutations for which these vectors add to the desired values.
The purpose of this work was to determine to what extent this approach is feasible when applied to structural changes in a protein in addition to the functional effects of mutations. If the additivity of effects of mutations on both structure and function is often very good and we are able to define conditions where it is and is not applicable, then our technique for specifically engineering individual properties of a protein may have very general applicability. It could be used to engineer proteins with altered substrate specificity but without decreased stability, for example, or to engineer proteins with increased catalytic activity but without altered substrate specificity.

**Gene V protein as a model system for protein engineering**

The model system for this study is the gene V protein of bacteriophage f1 (Figure 1). Gene V protein is a small, 87-residue, protein that binds tightly and cooperatively to single-stranded DNA and RNA (Alberts et al., 1972; Baas, 1985; Cavalieri et al., 1976; Pratt and Ehrdahl, 1968). Its natural function is to bind to a single-stranded DNA produced during infection of *E. coli* by bacteriophage f1, preventing its conversion into double-stranded DNA and instead allowing its packaging into new phage particles (Baas, 1985). The gene V protein is a dimer and we have recently determined its crystal structure at high resolution (Skinner et al., 1994). This structure shows which residues in the dimeric protein are close together and which are not, and hence are useful in our interpretation of reasons for deviations from additivity in effects of mutations on protein stability. In addition to the structure of the protein we have recently determined, we have previously isolated 350 single-amino-acid substitution mutants of the gene and characterized their function *in vivo* (Terwilliger et al., 1994; Zabin et al., 1991).

**Structural basis for additivity of mutational effects on the properties of a protein**
We have purified and characterized the stabilities and DNA-binding affinities of a collection of single and double mutants of gene V protein in which many of the mutations were at two buried sites, Val35 and Ile47 (Sandberg et al., 1995; Sandberg and Terwilliger, 1991; Sandberg and Terwilliger, 1993; Terwilliger, 1993). Mutations at these sites result in closely additive effects on the properties of the resulting proteins and we intended to determine whether there was a clear structural basis for this additivity. We were able to obtain high-quality crystals of 9 mutants of gene V protein with single and double amino acid substitutions at positions Val35 and Ile47 (Skinner and Terwilliger, 1996; Zhang et al., 1995). At position 35 the single mutant V35I was obtained, and at position 47, we obtained structures of the single mutants I47V, I47L, and I47M. The double mutants obtained were V35I/I47V, V35I/I47L, V35I/I47M, V35I/I47F, and V35A/I47L. Diffraction data of excellent quality (Rsym of approximately 4% on F for each dataset) were collected to a resolution of 1.8 Å from each of these crystals, all of which were closely isomorphous to crystals of wild-type gene V protein (space group C2).

Analysis of \((F_{\text{mut}} - F_{\text{wt}})e^{i\phi}c\) and \((2F_{\text{mut}} - F_{\text{wt}})e^{i\phi}c\) difference Fourier maps (obtained using phases calculated from the wild-type gene V protein model) it was clear that the structural changes caused by these mutations could be readily identified even though these structural shifts were typically on the order of 0.1 to 0.2 Å. Figure 2 illustrates a small part of the \((F_{\text{mut}} - F_{\text{wt}})e^{i\phi}c\) difference Fourier for the V35A/I47L double mutant in the neighborhood of position 35. The refined models for the wild-type and mutant structures are shown. The shifts in position of atoms near the V35A mutation are clearly indicated by the pattern of positive and negative difference density on either side of the atoms.

Additivity of structural effects of mutations
Upon examining the difference Fourier discussed above and the refined structures of single and double mutants of gene V protein, it was clear that structural changes in gene V protein caused by mutations at positions 35 and 47 are remarkably additive (Skinner and Terwilliger, 1996; Zhang et al., 1995). There are two ways in which this additivity is exhibited. The first way is illustrated in Figure 3. This figure shows the coordinate shifts in the side chain of Leu 81, a residue that is mid-way between residues 35 and 47 and that is affected by both the V35I and I47L mutations. The coordinate shifts in this side chain caused by the two mutations are very nearly additive in the double mutant.

A second reason for additivity of structural effects is that the regions affected by the two mutations are at least partially distinct (Skinner and Terwilliger, 1996; Zhang et al., 1995). In this case, one mutation affects one region, the other affects another, and therefore it is not surprising that the effects are additive. This is also the case for gene V protein mutants at positions 35 and 47. Fig. 4 shows positive and negative peaks of difference electron density which indicate the shifts in atomic positions that occur in the V35I single mutant. The peaks of difference density are clustered near the corresponding sites of mutation, with few structural changes extending far from the site. The regions affected by each mutation are not spheres about the mutated sites. Instead the effects are highly non-isotropic, evidently depending on which regions of the protein molecule have to adjust in order to accommodate the newly introduced mutated side chain. This is true in general for the gene V protein mutant structures we have examined (Skinner and Terwilliger, 1996; Zhang et al., 1995). Structural changes are relatively local, but are non-isotropically distributed about the site of mutation.

*Using the region structurally affected by single mutations to identify pairs of mutations that will show additive effects*
The local and highly non-isotropic nature of structural effects caused by mutations in the core of gene V protein suggested a simple way to identify cases where structural and functional effects of multiple mutations are probably additive. Comparing the structure of each mutant with a single amino acid substitution with that of the wild-type protein, the region influenced structurally by that mutation can be identified, either qualitatively as in Fig. 4, or more quantitatively. Pairs of mutants in which the regions of influence do not substantially overlap can be expected to result in generally additive effects (as double mutants at positions 35 and 47 do in gene V protein). In locations in the protein where the regions of influence do overlap, the structural changes might (as in Fig. 3) or might not show additive effects, and the extent of non-additivity might depend on the extent of structural changes in the individual mutants in the region of overlap. Addressing these issues will be an important area of future research in this field.

Acknowledgments

The authors would like to thank the NIH and the LDRD program of Los Alamos National Laboratory for generous support.
Figure legends

Figure 1. Gene V protein dimer (Skinner et al., 1994). The gene V protein is a symmetric dimer. The 2-fold axis of symmetry is approximately perpendicular to the page at the center of the figure.

Figure 2. V35A/I47L double mutant difference Fourier. The relevant part of the wild-type gene V protein structure is shown in white, the mutant in gray. Positive difference density is shown in solid contours, negative in dashed contours; contours are drawn at +/- 2.5 σ. Coordinate shifts for selected atoms are shown.

Figure 3. Additivity of structural changes in single and double mutants of gene V protein. The position of the side chain of Leu 81 in WT, V35I, I47L, and V35I/I47L mutants is shown. Directions of coordinate shifts are indicated by the arrows.

Figure 4. Difference Fourier map for V35I mutant. The relevant part of the wild-type gene V protein structure is shown as a line drawing. Positive difference density is shown in solid contours, negative in dashed contours; contours are drawn at +/- 4 σ.
References


Val 45 CG2 (0.30 A shift)

Val 70 CG2 (0.22 A shift)

Val 35

Leu 37 CD2 (0.54 A shift)