Relationship between *in Vivo* Activity and *in Vitro* Measures of Function and Stability of a Protein[†]

Warren S. Sandberg,[‡] Petra M. Schlunk,[§,||] Hal B. Zabin,[§,⊥] and Thomas C. Terwilliger*,[▽]

Department of Biochemistry and Molecular Biology and Department of Medicine, The University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois 60637, and Life Sciences Division, Mail Stop M880, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Received March 6, 1995; Revised Manuscript Received July 6, 1995[⊗] .

ABSTRACT: The *in vivo* activities of mutant proteins are readily measured and can potentially be used to estimate changes in *in vitro* properties such as stability or function, but this connection has not been rigorously established. Gene V protein is a small protein produced by bacteriophage f1 that binds to single-stranded DNA and to RNA and for which fitness can be assayed both *in vivo* and *in vitro*. We have assembled a large number of temperature-sensitive mutants of the gene V protein of bacteriophage f1 and measured their ability to support phage growth and replication *in vivo*. We have also purified many of these mutant gene V proteins and measured their stabilities and ssDNA binding affinities *in vitro*. Mutations at surface residues frequently yielded temperature-sensitive mutants, but remarkably, no overall correlation between *in vivo* activity and *in vitro* measures of either stability or function was found for this group. Mutations at buried residues often lead to the temperature-sensitive phenotype. At buried sites temperature sensitivity was strongly correlated with *in vitro* stability changes, but not with *in vitro* ssDNA binding affinity. The implication of these observations for protein engineering efforts is that phenotypes conferred by amino acid substitutions at buried sites can be used to identify mutants whose stabilities fall into ranges of interest, while phenotypes of mutants with surface substitutions may be much less readily interpreted, even in the case of a single-stranded-DNA-binding protein.

The goal of protein engineering is to design proteins with rationally devised novel properties. Because our understanding of how a particular mutation at a given site will affect the properties of interest is incomplete at best, it is usually necessary to make the desired mutation and then characterize the mutant protein *in vitro*. This can lead to a time-consuming trial and error approach to protein engineering. One partial exception to this situation can occur when a collection of characterized mutants exists. It is possible by using such a collection to combine single mutants to create double or multiple mutants whose properties are often close to the sum of the constituent single mutants (Sandberg & Terwilliger, 1993). In circumstances lacking a large number of previously characterized mutants, it would be very helpful to have a rapid means of identifying mutants with desired properties without resorting to purification and characterization of each candidate protein. Methods based on the sequence of the wild-type protein and observed frequencies of amino acid exchange found in nature are of little utility in predicting the *in vivo* activity of mutants (Terwilliger *et al.*, 1994; Zabin *et al.*, 1991). Although it is not possible to predict the activity of mutant proteins from the types of the amino acid substitutions involved, it may be more reasonable to make predictions about the stability and function of mutant proteins based on the observed *in vivo* activities. In this case, activity is being used as a secondary reporter for stability and function, using the assumption that active proteins must be sufficiently stable and close enough to the wild-type structure to carry out their activities *in vivo* (Lim & Sauer, 1989).

Many mutant proteins differing from the wild-type protein by one or a few amino acids show strongly temperature sensitive (TS)[1] behavior in their function (Horowitz, 1950). These proteins are active at a low, permissive temperature and have more or less compromised activity at higher, nonpermissive temperatures. Because it is presumably the alteration of some property of the substituted protein that causes the TS phenotype of the host organism, the substituted proteins encoded by temperature-sensitive organisms can be referred to as TS also (Pakula & Sauer, 1989). The ability to "stress" the mutant proteins *in vivo* by requiring them to operate at high temperature provides a means to compare mutants to each other and to the wild type.

The TS phenotype could result from decreases in stability, solubility, or resistance to proteolysis, as well as reduced function (i.e., catalysis, DNA binding, etc). The TS phenotype may also result from failure to accumulate sufficient quantities of active protein at the nonpermissive temperature, either because of poor expression, failure of the polypeptide

[†] This work was supported by the National Institutes of Health and the LDRD program of Los Alamos National Laboratory.

* To whom correspondence should be addressed.

[‡] Department of Medicine, The University of Chicago.

[§] Department of Biochemistry and Molecular Biology, The University of Chicago.

[||] Present address: 932 S. Sedona Ln., Anaheim Hills, CA 92808.

[⊥] Present address: Max-Planck-Institut fuer molekulare Genetik, Ihnestrasse 73, 14195 Berlin, Germany.

[▽] Los Alamos National Laboratory.

[⊗] Abstract published in *Advance ACS Abstracts,* September 1, 1995.

[1] Abbreviations: GdnHCl, guanidine hydrochloride; MOPS, 3-(*N*-morpholino)propanesulfonic acid; ssDNA, single-stranded DNA; NaCl; sodium chloride; poly-dA, polydeoxyadenylic acid; CD, circular dichroism; WT, wild type; NOE, nuclear Overhauser effect; TS, temperature sensitive.

0006-2960/95/0434-11970$09.00/0

to fold, or aggregation (Horowitz, 1950; Mitraki *et al.*, 1993; Mitraki & King, 1989). Destabilization of the folded protein is frequently cited as a cause for the TS phenotype (Pakula *et al.*, 1989). For example, many TS mutations of T4 lysozyme are found at positions which are buried and which have low thermal factors in the crystallographic model of the wild-type protein (Alber *et al.*, 1987). Thermal denaturation studies on four of these TS mutants show destabilization ranging from 1.2 to 3.1 kcal/mol relative to the wild-type protein, suggesting that destabilization of the folded protein may lead to the TS phenotype (Becktel & Baase, 1987; Hawkes *et al.*, 1984). TS mutations of λ repressor and λ cro are also frequently found at positions inaccessible to solvent in the folded protein (Hecht *et al.*, 1983; Pakula *et al.*, 1986). Generally, TS mutations are found most frequently at sites having less than 20% solvent exposure in iso-1-cytochrome *c*, T4 lysozyme, staphylococcal nuclease, λ cro, and λ repressor, and this distribution has been interpreted to mean that TS mutations frequently affect protein stability (Pakula *et al.*, 1989).

Previously we have studied multiple *in vitro* properties of 13 temperature-sensitive mutants of the bacteriophage f1 gene V protein (Zabin & Terwilliger, 1991). The gene V protein is an 87 amino acid homodimer that binds cooperatively to single-stranded DNA (ssDNA) (Alma *et al.*, 1983; Bulsink *et al.*, 1985), and active gene V protein is required for propagation of the phage (Alberts *et al.*, 1972; Pratt & Ehrdahl, 1968). The gene V protein has many activities *in vivo*. The primary function of the gene V protein is to bind to ssDNA, regulating the switch from the double-stranded replicative form to the production of a single-stranded form destined for extrusion through the cell membrane. In phage-infected cells, the gene V protein regulates translation of genes I, II, III, V, and X by binding to the 5′ untranslated region of their transcripts, although it has been suggested that this function is dispensable *in vivo* at least in the case of gene II message (Michel & Zinder, 1989; Yen & Webster, 1982; Zaman *et al.*, 1990, 1991, 1992). The binding of gene V protein to nucleic acids is highly cooperative *in vitro* (Alma *et al.*, 1983; Bulsink *et al.*, 1985; Stassen *et al.*, 1992).

TS mutants of the gene V protein differ from the WT in their DNA binding affinities, resistance to chemical denaturation, and rates of irreversible aggregation, suggesting that alterations of multiple properties occur in TS proteins and that all may contribute to the TS phenotype (Zabin *et al.*, 1991). Given the small number of TS gene V protein mutants characterized to date, it is still possible that a dominant factor leading to the TS phenotype could be identified by studying a larger number of mutants. Our approach to this question was to compare the *in vivo* phenotypes of a large collection of bacteriophage f1 gene V protein mutants with their corresponding *in vitro* DNA binding affinities and stabilities, after grouping mutations into surface and buried sites. We expected any general correlations of this type would be very useful in rapid *in vivo* screening of engineered mutants to ascertain whether they are likely to be functional or stable *in vitro*.

MATERIALS AND METHODS

Plasmids, Phage, and Strains. Mutagenesis of gene V was carried out in the plasmid pTT18 as described (Terwilliger, 1988a,b). The gene V *amber* phage f1T2 has been described (Terwilliger *et al.*, 1988). Single mutants were constructed by oligonucleotide-directed mutagenesis and isolated as derivatives of pTT18. Double mutants were obtained by oligonucleotide-directed mutagenesis, recombination of single mutants by the use of intervening restriction sites, or selection (as intragenic suppressers of a starting TS mutation) from a pool of random single amino acid substitution mutants (Zabin *et al.*, 1991). Mutant genes were expressed in *Escherichia coli* strain K561 (Davis *et al.*, 1985). Transformation of K561 with pTT18 derivatives was effected using an electroporation device.

Measurement of Phage Activity. Genes encoding mutant proteins were excised from derivatives of plasmid pTT18 and ligated in place of the null gene V in phage f1T2 (Terwilliger *et al.*, 1988). Restriction endonucleases *Mlu*I and *Kpn*I, cleaving pTT18 and codons 15 and 84, were used for the excision. Recombinant phage were tested for the ability to form plaques on lawns of *E. coli* at 34 and 40.5 °C as described earlier (Zabin *et al.*, 1991). Due to the high efficiency of transformation with the electroporation device used, the assay was readily reproducible, and at least 50 plaques were obtained at each temperature in all control experiments beginning with wild-type gene V in plasmid pTT18. Recombinant phage were classified according to their abilities to form plaques at 34 and 40.5 °C. Those that formed no plaques at either temperature were considered inactive (classified as ts-4). Those forming plaques similar in size to the wild type at both temperatures were considered fully active (classified as ts-0). Phage forming plaques similar in size to the wild type at 34 °C and slightly smaller than the wild-type phage at 40.5 °C were considered weakly TS (classified as ts-1). Those forming plaques similar to the wild type at 34 °C and much smaller than the wild type at 40.5 °C were considered TS (classified as ts-2). Phage that formed plaques similar to or somewhat smaller than those formed by the wild-type at 34 °C and formed no plaques at all at 40.5 °C were considered strongly TS (classified as ts-3). The phenotype of mutants were assessed in one to six independent trials, and the phenotypes of the mutants were quite reproducible. Using our arbitrary scale of TS values from 0 to 4, the mean value of the standard deviations for 27 mutants with four or more observations was just 0.35 unit. Two phage, L49M and L49F, formed plaques that were smaller than the WT at both temperatures and are not considered further here, as their phenotypes cannot be compared with the others in this study. Phage ssDNA from all active phage was isolated and sequenced using the dideoxy method to verify that the recombinant gene V contained the expected mutation. In cases where the phenotypes of these phage have been reported previously, the data reported here were taken from the original works (Terwilliger *et al.*, 1994; Zabin *et al.*, 1991). Fifteen of these phage have been tested for temperature sensitivity under slightly different conditions (34 and 42 °C) (Zabin *et al.*, 1991).

Protein Purification. Growth of K561 cultures transformed with pTT18 derivatives encoding gene V protein variants and purification of proteins were carried out as described (Liang & Terwilliger, 1991; Zabin *et al.*, 1991). To confirm that the plasmids encoding mutant proteins contained the expected nucleotide substitutions, single-stranded plasmid DNA was isolated from *E. coli* harvested late in the growth and the gene V region was sequenced using the dideoxy method (Sanger *et al.*, 1977).

Measurement of ssDNA Binding Affinity. NaCl-induced dissociation of gene V protein−ssDNA complexes, monitored by fluorescence, was used to estimate the binding affinities of WT and mutant gene V proteins for the substrate polydeoxyadenylic acid (poly-dA) as described (Alma *et al.*, 1983; Bulsink *et al.*, 1985; de Jong *et al.*, 1987; Zabin *et al.*, 1991). Data are reported as the apparent free energy change upon dissociation in 0.15 M NaCl ($\Delta G°_{d,0.15M}$), related to the effective binding constant ($K\omega_{bind,0.15M}$) by $\Delta G°_{d,0.15M} = RT \ln(K\omega_{bind,0.15M})$ [$R = 1.987$ cal/(mol °K), $T = 298$ K]. Dissociation is unfavorable at 0.15 M NaCl, so $\Delta G°_{d,0.15M}$ is positive. Differences in binding between mutants are expressed as differences in free energy change upon dissociation ($\Delta\Delta G°_{d,0.15M}$), defined as ($\Delta G°_{d,0.15M,mutant} - \Delta G°_{d,0.15M,WT}$). Mutants binding less tightly to ssDNA than the wild type will have negative values of $\Delta\Delta G°_{d,0.15M}$. Error estimates (2SD) were obtained from seven independent measurements of the DNA binding affinity of the WT gene V protein leading to errors of ±0.09 kcal/mol for $\Delta G°_{d,0.15M}$ and ±0.12 kcal/mol for $\Delta\Delta G°_{d,0.15M}$. In cases where the DNA binding affinities of these proteins have been reported previously, the data reported here were taken from the original works (Sandberg *et al.*, 1993; Zabin *et al.*, 1991).

Measurements of Protein Stability. Stability measurements on mutant gene V proteins at 25 °C were carried out as described (Liang *et al.*, 1991). The gene V protein is reversibly denatured by guanidine hydrochloride (GdnHCl), and the denaturation can be monitored by the disappearance of a tyrosine CD signal at 229 nm. Unfolding data were fitted to a two-state model (Liang *et al.*, 1991) with modifications (Sandberg & Terwilliger, 1991) in the case of proteins for which the unfolding is more than 50% complete when [GdnHCl] < 1.5 M. Stabilities are expressed as free energy changes upon unfolding, in kcal/mol of dimeric protein. Stabilities ($\Delta G°_{u,2M}$) are compared in the presence of 2.0 M GdnHCl, a denaturant concentration at which the unfolding equilibrium constants of most mutants can be compared directly, to minimize errors introduced by extrapolation to 0 M GdnHCl. The stability of the WT gene V protein in 2 M GdnHCl, given as the average ±2SD of 16 measurements, is 9.11 ± 0.47 kcal/mol. C_m is the molar concentration of denaturant required to bring the denaturation halfway to completion, and m is the dependence of $\Delta G°_u$ on [GdnHCl], with values (±2SD) of $C_m = 2.62 \pm 0.07$ M and $m = -3.69 \pm 0.58$ kcal/(mol M), respectively, based on 16 measurements of WT stability. The error estimate (2SD) for $\Delta\Delta G°_{u,2M}$, the value of stability changes between proteins, is ±0.67 kcal/mol. In cases where the stabilities of these proteins have been reported previously, the data reported here were taken from the original works (Sandberg & Terwilliger, 1989; Sandberg *et al.*, 1991, 1993; Zabin *et al.*, 1991).

RESULTS

Phenotypes, Stabilities, and DNA Binding Affinities of Mutants with Interior Amino Acid Substitutions. Table 1 gives the phenotypes, stability changes, and DNA binding affinity changes, relative to WT, for 68 single and double mutants at buried sites in the gene V protein. Positions having <10% solvent exposure of the WT side chain in the crystal structure are classified as buried (Skinner *et al.*, 1994). Mutants at positions 6, 33, 35, 37, 45, 47, 49, 63, 67, 68, 78, and 81, all having <10% exposure of the corresponding residue in the WT protein, are included in Table 1. In most cases, these mutants were isolated for the purpose of studying

Table 1: Mutations at Buried Positions[a]

mutations		solvent exposure			stability, $\Delta\Delta G°_{u,2M}$	DNA binding, $\Delta\Delta G°_{d,0.15M}$
position 1	position 2	position 1 (%)	position 2 (%)	TS	(kcal/mol)	(kcal/mol)
16V		0.0		0	−0.7	0.2
C33A		0.0		0	−0.5	
C33I		0.0		4	−0.9	
C33L		0.0		4	−2.6	
C33M		0.0		4	−3.5	−0.4
C33S		0.0		3	−4.3	0.0
C33T		0.0		4	−4.6	−0.4
C33V		0.0		1	−0.2	−0.1
V35A		0.0		0	−2.3	−0.5
V35C		0.0		0	−1.5	−0.6
V35F		0.0		3	−3.2	−1.0
V35I		0.0		0	−0.7	−0.2
V35L		0.0		3	−2.7	−0.5
V35M		0.0		0	−1.1	−1.0
V35T		0.0		3	−5.3	−0.5
L37A		9.8		4	−7.7	
L37C		9.8		4	−4.6	−0.8
L37I		9.8		0	−1.4	−0.4
L37T		9.8		3	−5.2	−0.1
L37V		9.8		2	−3.5	−0.4
V45A		0.0		1	−2.1	−0.6
V45C		0.0		3	−0.0	0.2
V45L		0.0		3	−3.0	
V45T		0.0		3	−3.5	−0.2
I47A		0.0		3	−7.1	0.2
I47C		0.0		2	−5.3	−0.3
I47F		0.0		1	−2.0	0.9
I47L		0.0		0	−0.7	0.0
I47M		0.0		1	−2.2	0.6
I47T		0.0		2	−7.6	1.2
I47V		0.0		0	−2.6	−0.2
L49A		0.0		4	−6.1	
L49C		0.0		4	−4.1	−1.0
L49I		0.0		2	−1.9	
L49T		0.0		4	−5.7	−1.0
L49V		0.0		2	−2.9	−0.8
V63C		3.4		2	−4.1	−0.1
V63T		3.4		3	−5.0	−0.3
S67C		0.0		4	−3.7	
S67T		0.0		1	−1.6	−0.3
F68L		3.1		3	−4.3	−0.3
F68V		3.1		4	−5.0	−3.1
I78C		6.2		3	−4.4	0.0
I78T		6.2		4	−6.6	−0.3
I78V		6.2		2	−1.3	−0.3
L81C		2.7		3	−3.7	−0.7
L81T		2.7		3	−5.1	−0.8
L81V		2.7		1	−0.2	−0.4
C33M	I47C	0.0	0.0	2	−5.7	0.3
C33V	V35C	0.0	0.0	1	−1.6	−0.6
V35A	I47A	0.0	0.0	3	−10.8	
V35A	I47F	0.0	0.0	1	−3.7	0.1
V35A	I47L	0.0	0.0	0	−3.0	−0.5
V35A	I47M	0.0	0.0	1	−4.5	0.1
V35A	I47V	0.0	0.0	1	−4.5	−0.9
V35C	I47C	0.0	0.0	3	−7.2	−0.8
V35F	I47L	0.0	0.0	3	−4.2	−1.3
V35I	I47F	0.0	0.0	1	−2.1	0.3
V35I	I47L	0.0	0.0	0	−1.2	−0.3
V35I	I47M	0.0	0.0	1	−2.8	0.5
V35I	I47V	0.0	0.0	2	−3.1	−0.5
V35L	I47F	0.0	0.0	2	−4.0	0.2
V35L	I47L	0.0	0.0	3	−3.6	−0.5
V35L	I47M	0.0	0.0	3	−5.4	−0.0
V35L	I47V	0.0	0.0	3	−5.1	−1.0
V35M	I47F	0.0	0.0	1	−2.4	0.0
V35M	I47L	0.0	0.0	1	−1.7	−1.1
V35M	I47M	0.0	0.0	1	−3.6	−0.5

[a] Substitutions and solvent exposures of WT residues at buried positions in the gene V protein are given. Temperature-sensitive phenotypes are categorized using a scale ranging from 0 (wild-type activity) to 4 (inactive) as described in the text. Changes in stability, measured as the free energy of unfolding, $\Delta\Delta G°_{u,2M}$, are given in kcal/mol of dimeric protein, relative to WT. Mutants with increased stabilities have positive values of $\Delta\Delta G°_{u,2M}$, and the magnitude corresponds to making each substitution twice. Changes in apparent free energies of dissociation from polydeoxyadenylic acid ($\Delta\Delta G°_{d,0.15M}$), relative to the WT protein, are given in kcal/mol; positive values of $\Delta\Delta G°_{d,0.15M}$ indicate enhanced binding of the mutant to ssDNA relative to WT.

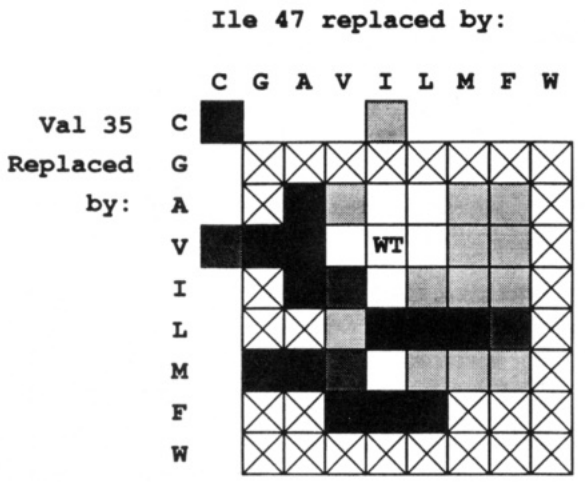

FIGURE 1: Activities of gene V protein mutants containing one or two substitutions in the hydrophobic core. Substitutions at position 47 are shown across the top, while substitutions at position 35 are shown down the side. Combinations of residues tested are boxed, with decreasing activity indicated by darker shading as described in Materials and Methods. Mutants with wild-type activity have no shading, while those that are severely TS are shaded black. Mutants that are inactive at either growth temperature are indicated with a ×.

the effects of interior apolar substitutions on protein stability, with assessment of phenotype being an incidental part of this process. The three mutants C33M,[2] F68L, and F68V were selected for the TS phenotype, and the mutant V45C was selected as a second-site suppressor of the F68L mutation. Other than this, the mutants in this set were obtained without selection. Because most of these mutants were selected neither for activity nor for the TS phenotype, this collection can be used to estimate the incidence of the TS phenotype among mutants at buried sites.

Roughly 85% of the mutants at buried apolar positions listed in Table 1 are temperature sensitive. This finding agrees with previous observations that substitutions at interior sites frequently cause the TS phenotype (Bowie & Sauer, 1989; Loeb *et al.*, 1989; Pakula *et al.*, 1989; Reidhaar-Olson & Sauer, 1990; Terwilliger *et al.*, 1994). The collection of mutants presented in Table 1 includes all of the apolar substitutions naturally possible at positions 35 and 47. These positions have been characterized as relatively tolerant to substitutions in the sense that functional proteins possessing apolar substitutions can be obtained (Terwilliger *et al.*, 1994). Both sites are completely shielded from solvent in the crystallographic structure of the gene V protein (Skinner *et al.*, 1994). Val35 and Ile47 are contained within the hydrophobic core but do not contact each other directly.

We analyzed the incidence of the TS phenotype in proteins containing multiple apolar interior substitutions by characterizing the activities of mutants containing all possible pairwise combinations of eight apolar amino acids at positions 35 and 47. Figure 1 summarizes graphically the phenotypes of a collection of mutants containing all of the pairwise combinations of Gly, Ala, Val, Ile, Leu, Met, Phe, and Trp at positions 35 and 47, as well as mutants containing Cys at one or both sites. Each box in Figure 1 denotes a mutant, while the interior shading of each box indicates the severity of the TS phenotype. Combinations of residues not enclosed by boxes were not tested. Mutants with activities indistinguishable from the wild type have no shading. Mutants considered weakly TS (classified as ts-1) are lightly shaded, while those considered TS (classified as ts-2) have heavy shading. Mutants considered strongly TS (classified as ts-3) are shaded black, and the mutants that do not support phage growth at even the permissive temperature (ts-4) are indicated with a ×. Of the 64 possible combinations (excluding Cys mutants) of apolar residues, 29 combinations lead to proteins that are not active *in vivo*. Of the combinations leading to active proteins, again excluding Cys mutants, 28 out of 35 are TS.

At both positions 35 and 47, mutations are tolerated in the sense that most proteins containing a mutation at just one of these sites support phage growth (Terwilliger *et al.*, 1994). However, as differences between the structures of the WT and substituting residues are increased, the activity of the resulting protein decreases. At position 35, Ala, Ile, or Met yields a protein that supports phage growth as well as the protein containing a WT Val. These residues are more similar to Val in terms of size than are Gly or Phe (which yield inactive or extremely TS mutants) and do not move the branch point of the side chain distally, as does Leu. At position 47, only Val and Leu lead to proteins that support phage growth as well as the WT Ile residue. These residues are similar to Ile in terms of size (in the case of Leu) and the maintenance of a β-branched structure (in the case of Val). Replacement of Ile at position 47 with Met yields a protein that is mildly TS, despite the similarity in size between these two residues.

The observation in Figure 1 that some single amino acid substitution mutants have wild-type activity raises the question of whether these mutations can be combined to give double mutants whose activities are also wild type. A related question is whether active double mutants can be made by combining inactive or TS single mutants at these two sites. Figure 1 shows that only one of the double mutants (V35A/I47L) is as active as the WT, and no double mutants were found whose activities were greater than the activities of their constituent double mutants. In fact, combinations of single mutants displaying wild-type activity (such as V35I combined with I47V to make V35I/I47V) lead to double mutants which are TS in three of the four cases shown in Figure 1. This result suggests that even conservative substitutions produce slight changes in phenotype which become evident when the mutations are combined. In general, combining two single mutants leads to a double mutant whose TS phenotype is equally severe or more so, indicating that the effects of mutations on gene V protein activity are cumulative, as suggested earlier (Sandberg *et al.*, 1993).

What is the basis for the TS phenotype frequently found in mutants at buried sites? Because side chains at buried positions are entirely shielded from solvent, they are expected to play little direct role in protein−ssDNA or protein−protein contacts. On the other hand, substitutions at these positions are likely to alter stability, both by changing the interior hydrophobicity and by their effects on the packing arrangement of interior residues (Eriksson *et al.*, 1992, 1993; Karpusas *et al.*, 1989; Kellis *et al.*, 1988, 1989; Lim *et al.*, 1992; Mendel *et al.*, 1992; Sandberg, *et al.*, 1989, 1991). Finally, interior apolar substitutions can cause subtle alterations in protein structure which, in turn, can be propagated over relatively long distances (Alber, 1989). Thus, interior substitutions can reasonably be expected to affect properties

[2] Substitutions are described in the one-letter code; for example, Y41F denotes the replacement of tyrosine at position 41 by phenylalanine.

Table 2: Mutations at Exposed Positions and Double Mutants at Exposed and Buried Positions[a]

mutations		solvent exposure			stability, $\Delta\Delta G°_{u,2M}$	DNA binding, $\Delta\Delta G°_{d,0.15M}$
position 1	position 2	position 1 (%)	position 2 (%)	TS	(kcal/mol)	(kcal/mol)
V19C		57.3		1	−0.3	−0.7
V19T		57.3		1	−0.6	−0.3
K24V		82.5		1	0.8	−1.8
Y26R		33.7		4	−0.4	−2.0
L28V		37.2		0	1.1	−0.0
E30F		17.9		3	2.0	2.3
E30M		17.9		3	0.6	1.9
E30N		17.9		3	−1.1	1.4
L32H		47.6		0	−0.9	
L32R		47.6		0	−1.6	1.7
L32W		47.6		0	2.8	
D36C		31.2		3	−2.1	−0.0
E40C		86.0		2	−1.6	−0.4
E40T		86.0		4	−0.4	−1.5
Y41A		83.0		4	−0.4	
Y41F		83.0		4	−0.6	−3.1
V43C		31.0		0	−2.1	−0.8
V43T		31.0		0	−1.6	−0.6
T48C		16.9		3	−0.8	−0.4
T48V		16.9		0	0.0	0.1
D50H		33.6		2	−1.6	0.7
T62C		33.6		2	−0.7	−0.2
T62V		33.6		0	1.3	−0.2
K69H		44.4		3	−1.3	
K69M		44.4		3	0.1	−3.2
V70C		24.3		3	−3.3	−0.6
V70P		24.3		3	−5.1	−0.7
V70T		24.3		2	−3.5	−0.6
F73W		82.4		0	0.8	1.2
M77A		45.2		0	−2.1	
M77C		45.2		2	0.0	−0.3
M77F		45.2		1	−0.2	
M77I		45.2		1	1.6	0.1
M77L		45.2		1	−1.2	
M77T		45.2		2	−0.8	−0.3
M77V		45.2		0	1.2	0.0
R82C		25.2		3	−1.5	−2.7
L32Y	R82C	47.6	25.2	0	0.1	−1.8
I6V	M77I	0.0	45.2	0	0.3	0.2
I6V	M77V	0.0	45.2	1	−0.0	0.3
L28V	F68L	37.2	3.1	0	−3.5	−0.8
V45C	R82C	0.0	25.2	0	−1.0	−2.4
H64C	F68L	41.1	3.1	0	−4.1	−0.8
L65P	F68L	33.9	3.1	0	−4.3	−0.8

[a] Substitutions and solvent exposures of WT residues at exposed positions in the gene V protein are given. Temperature-sensitive phenotypes, changes in stability, and changes in DNA binding affinity are given as described in Table 1.

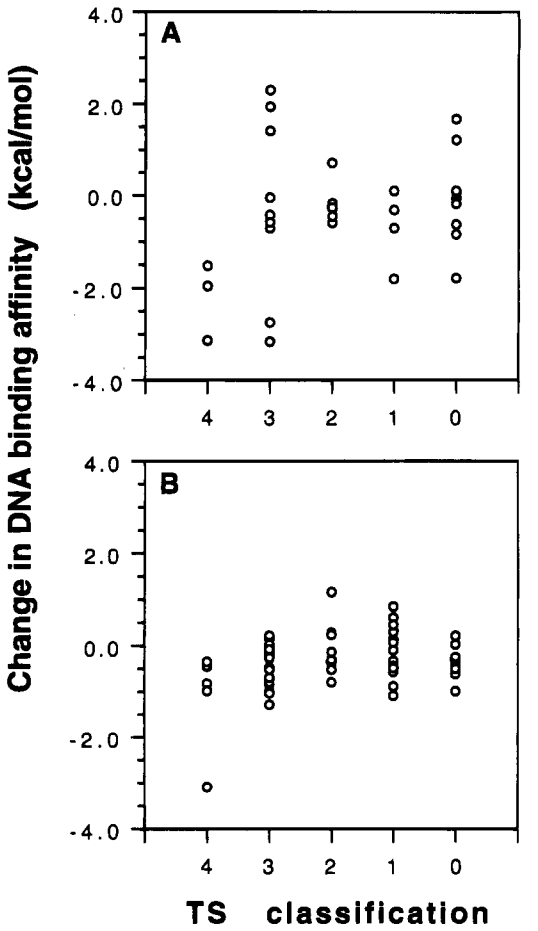

FIGURE 2: Changes in DNA binding affinity compared with TS phenotypes for mutants at exposed (A) and buried positions (B). Binding affinity changes ($\Delta\Delta G°_{d,0.15M}$), relative to the WT protein, are shown on the y axis. The TS classifications of the mutants are given on the x axis. TS classifications are determined as described in Materials and Methods, with TS = 0 indicating mutants with wild-type activity and TS = 4 indicating mutants that are inactive. A positive value of $\Delta\Delta G°_{d,0.15M}$ indicates enhanced binding to ssDNA relative to WT. The mean (±SEM) changes in DNA binding affinity as a function of TS classification for mutants at exposed positions were −0.0 ± 0.3, −0.7 ± 0.4, −0.2 ± 0.2, −0.2 ± 0.6, and −2.2 ± 0.5 kcal/mol for TS values of 0, 1, 2, 3, and 4, respectively. For mutants at buried sites the corresponding mean changes in DNA binding affinity were −0.3 ± 0.1, −0.1 ± 0.1, −0.1 ± 0.2, −0.4 ± 0.1, and −1.0 ± 0.4 kcal/mol, respectively.

such as DNA binding affinity by causing slight alterations in the arrangement of the residues involved in forming the protein−ssDNA complex. Consistent with these expectations, most of the mutants in Table 1 show changes in both stability and DNA binding affinity.

Phenotypes, Stabilities, and DNA Binding Affinities of Mutants with Surface Substitutions. Table 2 gives the phenotypes, stability changes, and DNA binding affinity changes, relative to WT, for 44 single and double mutants at exposed sites in the gene V protein, including five double mutants with one mutation at an exposed site and one at a buried site. Positions having >10% solvent exposure of the WT side chain in the crystal structure are classified as exposed (Skinner *et al.*, 1994). Most of the mutants in Table 2 were obtained by selection for the TS phenotype or as intragenic suppressors of other TS mutations. A simplistic expectation arising from this fact and the surface location of these sites is that most of the mutants in Table 2 would have as their major difference from the WT protein an alteration in DNA binding affinity or in protein−protein interactions, with stability being relatively unchanged. In fact, however, most of the mutants listed in Table 2 have alterations in both DNA binding and stability, suggesting that both factors affect the activities of mutants with amino acid substitutions at surface sites. This observation is consistent with the recent analysis of stabilities and structures of mutants in the active site of phage T4 lysozyme (Shoichet *et al.*, 1995). In this study most proteins at surface locations in the active site of T4 lysozyme that affected catalysis also substantially affected the stability of the protein, generally increasing the stability of the protein.

Relationship between Phenotype and DNA Binding Affinity. We have assessed the relationship between phenotype and *in vitro* measures of gene V protein function such as DNA binding affinity by plotting DNA binding affinity changes as a function of phenotype. In Figure 2, panel A shows changes in DNA binding affinity, relative to the WT, plotted according to TS classification for mutations at

exposed sites, while panel B shows the same information for mutants at buried sites. The five double mutants listed in Table 1 with one substitution at a buried site and one at an exposed site are not included in Figures 2 and 3, as they could be included in either category. The TS classification given along the bottom of Figures 2 and 3 is a measure of the severity of the activity decrement caused by amino acid substitutions. The consistent spacing between entries along the bottom of Figure 2A is not meant to imply that the phenotype classification can be treated as a continuous variable. However, the TS classification does establish a rank order of activity changes between groups. Changes in DNA binding affinity are reported on the vertical axis. DNA binding affinities are reported as the free energy change upon dissociation ($\Delta G°_{d,0.15M}$), which is positive. Therefore, mutants binding less tightly to DNA have negative values of $\Delta\Delta G°_{d,0.15M}$ relative to the wild-type protein.

Mutations at exposed sites, shown in Figure 2A, are associated with a wide range of changes in DNA binding affinity. In contrast, Figure 2B shows that although mutations at buried sites do affect DNA binding, as a whole the range of changes in DNA binding affinity is smaller within each TS category. The single inactive mutant (F68V) with markedly reduced binding affinity shown in Figure 2B shows that interior substitutions can at least in some cases dramatically affect ssDNA binding affinity of the gene V protein.

Relationship between Phenotype and Stability Changes in Mutant Proteins. In a procedure analogous to that used above, we assessed the association between phenotype and the stability of the mutant proteins by plotting $\Delta\Delta G°_{u,2M}$, relative to the WT, as a function of the TS classification. In Figure 3, panel A shows changes in stability, relative to the WT, plotted according to TS classification for mutations at exposed sites, while panel B shows the same information for mutants at buried sites. Again, the TS classification given along the bottom of the Figure 3 is a rank order measure of the severity of the activity decrement caused by amino acid substitutions. Changes in stability ($\Delta\Delta G°_{u,2M}$) are reported on the vertical axis. Stability is measured as the free energy change upon unfolding ($\Delta G°_{u,2M}$) and is therefore generally positive. When proteins are compared, the *difference* in free energy changes upon unfolding between the wild type and a destabilized mutant ($\Delta\Delta G°_{u,2M}$) will be negative.

Figure 3A shows that there is little apparent association between stability changes and TS category for mutations occurring at exposed sites. However, Figure 3B does show a strong association for mutations occurring at buried sites. In other words, mutants at buried sites manifesting temperature-sensitive (ts-2), strongly temperature sensitive (ts-3), or inactive (ts-4) phenotypes have lower stabilities, as a group, than those mutants with wild-type phenotypes.

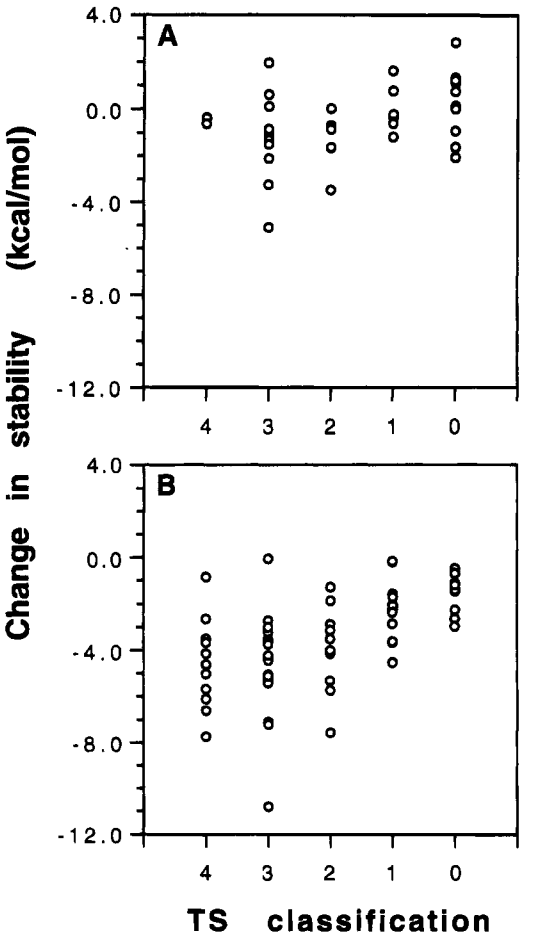

FIGURE 3: Changes in stability compared with TS phenotypes for mutants at exposed (A) and buried positions (B). Stability changes ($\Delta\Delta G°_{u,2M}$), relative to the WT protein, are shown on the *y* axis. The TS classifications of the mutants are given on the *x* axis. TS classifications are determined as described in Materials and Methods, with TS = 0 indicating mutants with wild-type activity and TS = 4 indicating mutants that are inactive. A positive value of $\Delta\Delta G°_{u,2M}$ indicates enhanced stability relative to WT. The mean (±SEM) changes in stability for mutants at exposed positions were -0.1 ± 0.5, 0.0 ± 0.4, -1.4 ± 0.5, -1.2 ± 0.6, and -0.4 ± 0.1 kcal/mol for TS values of 0, 1, 2, 3, and 4, respectively. For mutants at buried sites the corresponding mean changes in stability were -1.4 ± 0.3, -2.3 ± 0.3, -3.9 ± 0.6, -4.7 ± 0.5, and -4.6 ± 0.5 kcal/mol for TS values of 0, 1, 2, 3, and 4, respectively.

DISCUSSION

The purpose of this work was to examine whether the effects of a mutation on the properties of a protein *in vivo*, as monitored by an expressed phenotype such as temperature sensitivity, could be used to predict the effects of that mutation on the properties of the protein *in vitro*. A general correlation between *in vivo* and *in vitro* properties would simplify the task of engineering proteins because many mutants could be screened without the need to purify the proteins. Our approach was to examine the temperature sensitivities, stabilities, and DNA binding affinities of a collection of 112 mutants of the gene V protein of the bacteriophage f1. Mutations were divided into those occurring at the surface and those occurring at buried residues (Skinner *et al.*, 1994) because it was anticipated that the effects of mutations would be different in these two groups. TS or inactive mutations are generally thought to occur at buried sites, for example, while mutations at surface sites are frequently tolerated (Alber *et al.*, 1987; Pakula *et al.*, 1989; Reidhaar-Olson, *et al.*, 1990; Terwilliger *et al.*, 1994).

Much evidence has accumulated indicating that interior substitutions frequently lead to proteins with reduced activity (Alber *et al.*, 1987; Pakula *et al.*, 1989; Reidhaar-Olson *et al.*, 1990; Terwilliger *et al.*, 1994). Systematic examination of the incidence of the TS phenotype at positions 35 and 47 shows that roughly half of the side-chain pairs tested lead to active proteins (Figure 1). Most of the active mutants at these positions are TS, with wild-type activity found in only 7 of 64 possible pairwise combinations of eight apolar side chains. The finding that half of the mutants are active indicates that proteins commonly tolerate interior substitutions without complete loss of function. However, the

frequent occurrence of temperature sensitivity implies that only a few combinations of residues at positions 35 and 47 are tolerated with wild-type activity. In general, larger departures from the size and shape of the wild-type residue increase the likelihood that an interior substitution mutant of the gene V protein will be TS or inactive. For example, side chains such as Gly and Trp at positions 35 and 47 generally do not support phage growth (Figure 1). In another example of this property, the N-terminal domain of λ repressor can accommodate many combinations of interior substitutions and retain activity as long as the new side chains are generally nonpolar and the total volume of the hydrophobic core does not deviate significantly from the wild-type volume (Lim *et al.*, 1989). The high incidence of gene V protein interior substitution mutants that are active but temperature sensitive is at once consistent with the expectation that TS mutations frequently occur at buried sites and with the observation that interior sites can accommodate substitutions as long as the new residues are similar to the wild-type side chains. Presumably, small deviations from the wild-type residues perturb the stability or function of the protein less severely than mutations involving dissimilar residues. Crystallographic structures of the single mutants V35I and I47V and the double mutant V35I/I47V suggest that substitutions involving similar residues produce minor rearrangements of the rest of the protein hydrophobic core (unpublished observations). These rearrangements along with the changes in hydrophobicity found in the single mutants all decrease the stabilities of the mutant proteins (Table 1) (Sandberg *et al.*, 1989). Mutations such as I47A may introduce more drastic changes in the protein interior, which in turn destabilize the protein and render phage harboring these mutations TS (Table 1).

In the gene V protein, a group of 13 TS mutants distributed throughout the structure have been shown to possess alterations of stability, DNA binding affinity, and lifetime at elevated temperature, or combinations of these properties (Zabin, et al., 1991). On the basis of these results, the TS phenotype for a given mutant appears to result from defects in multiple properties. However, mutants at buried positions such as I47A, although severely destabilized, possess relatively normal DNA binding characteristics, while mutants at surface positions such as Y41F have relatively normal stability but bind poorly to ssDNA. Thus, mutants at buried positions might, as a group, have stability changes as their major defects, while mutants at exposed positions might have relatively larger defects in DNA binding. To test this, we compared the *in vivo* measurements of TS phenotype in the larger group of mutants considered here with the stability and DNA binding affinity changes of these mutants measured *in vitro*.

When the stability changes resulting from surface substitutions of the gene V protein were measured, there were no differences in the average stability changes found among groups of mutants having normal or decreased activity *in vivo* (Figure 3A). Mutations at surface sites are not commonly expected to exert strong effects on stability, although a "reverse" hydrophobic effect, destabilizing the folded protein, has been suggested to account for the destabilizations caused by apolar substitutions at highly exposed positions (Pakula & Sauer, 1990). Additionally, active site mutations frequently increase protein stability, presumably because they remove energetically unfavorable contacts (Shoichet *et al.*, 1995). Because some surface residues contact the substrate or are involved in protein–protein interactions, mutations at these sites could affect *in vitro* measures of function. This expectation is tempered by the fact that, depending on the relative sizes of the total protein surface and its functionally important surfaces, only a fraction of the surface residues are functionally important. The rest of the residues are expected to be relatively tolerant of substitution (Reidhaar-Olson *et al.*, 1990; Terwilliger *et al.*, 1994). Thus, in a random search for critical surface residues, a small active site means more surface residues are dispensable, decreasing the change of finding surface residues important for function.

In the case of the gene V protein, functionally important surface residues include those involved in ssDNA binding and those essential for protein–protein interactions leading to the observed cooperative nature of DNA binding. Five surface residues are very close to bound spin-labeled oligonucleotides in NMR experiments, while an additional 17 surface residues are somewhat more distant (de Jong *et al.*, 1989; Folkers *et al.*, 1993; van Duynhoven *et al.*, 1990, 1993). The residues closest to bound spin-label are Gly18, Ser20, Tyr26, Leu28, and Phe73. NOEs to bound oligonucleotides have been demonstrated for Tyr26, Leu28, and Phe73 (King & Coleman, 1987, 1988). Of the other 17 residues located near bound spin-label, five are relatively intolerant of substitution in the sense that amino acid replacements at these positions frequently impair the activity of the resulting proteins *in vivo* (Terwilliger *et al.*, 1994). These include Arg16, Val19, Arg21, Lys46, and Arg80, four of which are positively charged and could be expected to participate in binding polyanionic nucleic acids (Terwilliger *et al.*, 1994). Three other residues, Glu40, Tyr41, and Lys69, are thought to participate in cooperative dimer–dimer interactions (Folkers *et al.*, 1991; Skinner *et al.*, 1994; Terwilliger *et al.*, 1994). All of the aforementioned residues are at the protein surface, with solvent exposures ranging from 25% to 100%. Thus, at least 13 of the 87 amino acid residues in the gene V protein monomer appear to contribute to the functional surfaces.

Analysis of the DNA binding characteristics of gene V protein mutants bearing surface substitutions shows that mutants binding more strongly to ssDNA *in vitro* than the WT can be found in most TS categories (Figure 2A). Also, Figure 2A shows that a wide range of ssDNA binding affinities are encountered within each TS category. This may be due in part to the way that these mutants were isolated. Many of the proteins in the collection of mutations at surface sites were obtained by selection for the TS phenotype, while a few were the result of directed mutagenesis to produce substituted proteins for other experiments. Selection for the TS phenotype among mutants at surface residues is expected to yield a collection enriched for mutants with alterations in functionally important properties. However, the observed phenotypes could depend on multiple properties of the gene V protein, including proper folding and dimerization of the polypeptide; the solubility, stability, protease resistance, and lifetime of the folded protein; and the protein–protein and protein–ssDNA interactions necessary for normal function. Proteins with changes in many of these properties are likely to be obtained by selection for the TS phenotype among mutants at surface residues (Zabin *et al.*, 1991), and some of these are expected to have diminished ssDNA binding. Some of the proteins isolated as TS mutants such as E30N and E30F bind more strongly to ssDNA *in vitro* than does

the WT (Table 2; Figure 2A), a characteristic that was unexpected of a TS protein (Zabin *et al.*, 1991). Upon further study, these mutants were found to aggregate rapidly *in vitro*, suggesting that reduced intracellular lifetime may account for the temperature sensitivity *in vivo* (Zabin *et al.*, 1991). Selection for the TS phenotype apparently yields a collection of mutants altered in many properties, with mutants bearing predominant DNA binding affinity changes comprising a subpopulation of this group.

If altered ssDNA binding affinity were the predominant cause of the TS phenotype at surface sites, then we would expect that there would be a strong correlation between phenotype and changes in DNA binding. The wide variation in ssDNA binding affinities encountered in each group of TS mutants at exposed sites and the observation that strongly TS mutants can possess supranormal DNA binding affinity indicates that many other factors besides altered DNA binding affect the observed phenotypes. Affinity for ssDNA does not appear to have a dominant role in the determination of the resulting phenotype. The ssDNA binding affinity of a mutant falling into any TS classification, including those incapable of supporting phage growth, can approach that of the wild type. In light of these findings, *in vivo* phenotype is not a useful secondary reporter of ssDNA binding affinity in the case of mutations at exposed sites in the gene V protein.

Substitutions at buried positions also alter DNA binding affinity, with DNA binding affinities spanning about 2.5 kcal/mol for each group of TS mutants (Figure 2B). Interior substitutions in the gene V protein at positions 35 and 47 produce slight structural alterations (unpublished observations), and such structural changes accompanying substitutions at various buried sites could potentially be propagated to residues involved in DNA binding. This, in turn, presumably accounts for the changes in the DNA binding affinities of mutants at buried positions. However, the observed changes in DNA binding affinity encompass similar averages within each TS classification (Figure 2B). In other words, each group of mutants binds to ssDNA as well as the others, regardless of the severity of the TS phenotype. Thus the phenotypes observed *in vivo* appear to have little systematic dependence on the DNA binding affinities measured *in vitro*.

A possible exception to this generalization occurs in the case of mutants at buried sites leading to the inactive phenotype (Figure 2B). Of the six such mutants characterized, five had ssDNA binding affinities similar to that of the WT protein, while one, F68V, had a binding affinity 3.1 kcal/mol less than the WT (Figure 2B; Table 1). Because Phe68 is part of the dimer interface, replacement with a smaller residue may lead to a relatively large change in conformation and an associated reduction in DNA binding affinity because the arrangement of the DNA binding residues has been changed.

As the binding of gene V protein to single-stranded DNA requires dimerization of the protein, it is worth examining whether changes in the *in vitro* DNA binding affinity observed for mutants are likely to be due to changes in the dimerization state of the protein. The dimerization of wild-type gene V protein is closely coupled to the folding of the protein (Liang *et al.*, 1991), and stable, monomeric variants of the protein have not been reported. This suggests that the dissociation constant for a gene V protein variant dimer is roughly equal to the equilibrium constant for the unfolding reaction. For the wild-type gene V protein the equilibrium constant for unfolding is about 1.7×10^{-12} M at 25 °C in a buffered solution containing 0.1 M NaCl at pH 7 (Liang *et al.*, 1991), which means that, at the protein concentration of 2 μM used for the ssDNA binding experiments, about 99.9% of the protein is in the dimeric form. A variant which has a free energy of unfolding about 5 kcal/mol less than the wild-type protein would have an equilibrium constant of about 8×10^{-9} M under the same conditions, resulting in about 94% of the protein being in the dimeric form and 6% existing as unfolded monomers. This analysis indicates that, for gene V protein variants with free energies of unfolding within about 5 kcal/mol of that of the wild-type, partial unfolding of the protein and conversion into monomers would only have a small effect on the *in vitro* ssDNA binding affinity of the protein at 25 °C. At higher temperatures, such as those used for selection of the TS phenotype *in vivo*, a much higher fraction of each of the proteins, and particularly of the more unstable proteins, would be expected to be unfolded and therefore unable to contribute to binding ssDNA.

The final question to be answered concerns the association between stability changes and the observed phenotypes of mutants at buried positions. In the case of the gene V protein, Figure 3B demonstrates that the severity of the TS phenotypes measured *in vivo* is quite strongly associated with the stability changes measured *in vitro*. Proteins encoded by TS mutants of the gene V protein are, on average, less stable than those proteins encoded by genes yielding WT activity *in vivo*. For example, proteins with a TS classification of 3 (strongly temperature sensitive proteins) have average stability changes relative to the WT of −4.7 kcal/mol for 20 mutants, while the average stability change for those with a TS classification of 0 (WT phenotype) is just −1.4 kcal/mol for 11 mutants. In other proteins, such as T4 lysozyme and λ repressor, TS mutants at buried positions are frequently less stable than the WT, and this observation has led to the prediction that mutations at interior sites are generally TS because they destabilize the protein (Alber *et al.*, 1987; Hecht *et al.*, 1983). Using large numbers of gene V protein mutants bearing substitutions at buried positions, we have strengthened this argument by showing that TS mutants involving buried residues are less stable *as a group* than such mutants with wild-type phenotypes and that these differences are unlikely to have occurred by chance.

Comparison of Figure 3B with Figure 3A shows that mutants at buried sites tend to be less stable than mutants with corresponding phenotypes occurring at exposed positions. As the severity of the TS phenotype worsens, the average stability of mutants at buried positions decreases. For mutations at buried sites, the average destabilization, relative to the WT protein, increases from 1.4 to 4.6 kcal/mol as the phenotype progresses from wild type (ts-0) to strongly temperature sensitive (ts-4). The trend of increasing destabilization with worsening phenotype suggests that inactive proteins bearing interior substitutions may be drastically destabilized. Mutants of the gene V protein such as I47G and I47W are inactive and cannot be recovered by the standard purification procedure, although SDS−PAGE of crude extracts shows that cells expressing these mutant genes accumulate a polypeptide comigrating with the WT gene V protein (unpublished observations). Purification involves overexpression in *E. coli* followed by ssDNA affinity chromatography, suggesting either that polypeptides

with large changes in interior volume do not accumulate in the soluble, folded form *in vivo* or that they adopt a conformation that does not bind to ssDNA. Studies of the λ repressor show that mutants with interior substitutions leading to large changes in the volume of buried residues are found to be inactive and are highly unstable (Lim *et al.*, 1992). Structural analysis shows that these mutants are in a molten-globule state *in vitro*, rather than assuming the properly folded form (Lim *et al.*, 1992), and this situation may occur in the least stable of the inactive mutants of the gene V protein as well.

The association between temperature sensitivity observed *in vivo* and stabilities measured *in vitro* suggests that to a certain extent it is possible to predict the stability of a mutant on the basis of the temperature-sensitive phenotype conferred by the substitution for mutants at buried positions. However, Figure 3B shows that there is considerable overlap in the range of stability changes observed within each TS classification. The association between stability changes and activity changes for mutants at buried sites allows us to predict with some confidence that mutants at buried sites conferring the TS phenotype will be less stable than the wild type. For mutants at buried positions, the phenotype therefore becomes a useful secondary reporter of the likely stability of the mutant protein. For the purposes of protein engineering, this could shorten the search for mutants possessing desired interior substitutions without reductions in stability, because *in vivo* phenotype information is usually more readily obtained than are the stabilities of purified proteins.

ACKNOWLEDGMENT

We would like to thank H. Broihier, E. Manning, D. Cheng, and S. Choi for assistance in construction and characterization of mutants.

REFERENCES

Alber, T. (1989) *Annu. Rev. Biochem. 58*, 765−798.
Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C., & Matthews, B. W. (1987) *Biochemistry 26*, 3754−3758.
Alberts, B., Frey, L., & Delius, H. (1972) *J. Mol. Biol. 68*, 139−152.
Alma, N. C. M., Harmsen, B. J. M., de Jong, E. A. M., van der Ven, J., & Hilbers, C. W. (1983) *J. Mol. Biol. 163*, 47−62.
Becktel, W. J., & Baase, W. A. (1987) *Biopolymers 26*, 619−623.
Bowie, J. U., & Sauer, R. T. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2152−2156.
Bulsink, H., Harmsen, B. J. M., & Hilbers, C. W. (1985) *J. Biomol. Struct. Dyn. 3*, 227−247.
Davis, N. G., Boeke, J. D., & Model, P. (1985) *J. Mol. Biol. 181*, 111−121.
de Jong, E. A. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1987) *Biochemistry 26*, 2039−2046.
de Jong, E. A. M., van Duynhoven, J. P. M., Harmsen, B. J. M., Tesser, G. I., Konings, R. N. H., & Hilbers, C. W. (1989) *J. Mol. Biol. 206*, 133−152.
Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992) *Science 255*, 178−183.
Eriksson, A. E., Baase, W. A., & Matthews, B. W. (1993) *J. Mol. Biol. 229*, 747−769.
Folkers, P. J. M., Stassen, A. P. M., van Duynhoven, J. P. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1991) *Eur. J. Biochem. 200*, 139−148.
Folkers, P. J. M., van Duynhoven, J. P. M., van Lieshout, H. T. M., Harmsen, B. J. M., van Boom, J. H., Tesser, G. I., Konings, R. N. H., & Hilbers, C. W. (1993) *Biochemistry 32*, 9407−9416.
Hawkes, R., Grutter, M. G., & Schellman, J. (1984) *J. Mol. Biol. 175*, 195−212.
Hecht, M. H., Nelson, H. C. M., & Sauer, R. T. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2676−2680.
Horowitz, N. H. (1950) *Adv. Genet. 3*, 33−71.
Karpusas, M., Baase, W. A., Matsumura, M., & Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8237−8241.
Kellis, J. T., Jr., Nyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature 333*, 784−786.
Kellis, J. T., Jr., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914−4922.
King, G. C., & Coleman, J. E. (1987) *Biochemistry 26*, 2929−2937.
King, G. C., & Coleman, J. E. (1988) *Biochemistry 27*, 6947−6953.
Liang, H., & Terwilliger, T. C. (1991) *Biochemistry 30*, 2772−2782.
Lim, W. A., & Sauer, R. T. (1989) *Nature 339*, 31−36.
Lim, W. A., Farruggio, D. C., & Sauer, R. T. (1992) *Biochemistry 31*, 4324−4333.
Loeb, D. D., Swanstrom, R., Everitt, L., Manchester, M., Stamper, S. E., & Hutchison, C. A., III (1989) *Nature 340*, 397−400.
Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., & Schultz, P. G. (1992) *Science 256*, 1798−1802.
Michel, B., & Zinder, N. D. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4002−4006.
Mitraki, A., & King, J. (1989) *Bio/Technology 7*, 690−697.
Mitraki, A., Danner, M., King, J., & Seckler, R. (1993) *J. Biol. Chem. 268*, 20071−20075.
Pakula, A. A., & Sauer, R. T. (1989) *Annu. Rev. Genet. 23*, 289−310.
Pakula, A. A., & Sauer, R. T. (1990) *Nature 344*, 363−364.
Pakula, A. A., Young, V. B., & Sauer, R. T. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8829−8833.
Pratt, D., & Ehrdahl, W. S. (1968) *J. Mol. Biol. 37*, 181−200.
Reidhaar-Olson, J. F., & Sauer, R. T. (1990) *Proteins: Struct., Funct., Genet. 7*, 306−316.
Sandberg, W. S., & Terwilliger, T. C. (1989) *Science 245*, 54−57.
Sandberg, W. S., & Terwilliger, T. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1706−1710.
Sandberg, W. S., & Terwilliger, T. C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8367−8371.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Shoichet, B. K., Baase, W. A., Kuroki, R., & Matthews, B. W. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 452−456.
Skinner, M. M., Zhang, H., Leschnitzer, D. H., Bellamy, H., Sweet, R. M., Gray, C. M., Konings, R. N. H., Wang, A. H.-J., & Terwilliger, T. C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2017−2075.
Stassen, A. P. M., Harmsen, B. J. M., Schoenmakers, J. G. G., Hilbers, C. W., & Konings, R. N. H. (1992) *Eur. J. Biochem. 206*, 605−612.
Terwilliger, T. C. (1988a) *Gene 71*, 41−47.
Terwilliger, T. C. (1988b) *Gene 69*, 317−324.
Terwilliger, T. C., Fulford, W. D., & Zabin, H. B. (1988) *Nucleic Acids Res. 16*, 9027−9039.
Terwilliger, T. C., Zabin, H. B., Horvath, M. P., Sandberg, W. S., & Schlunk, P. M. (1994) *J. Mol. Biol. 236*, 556−571.
van Duynhoven, J. P. M., Folkers, P. J. M., Stassen, A. P. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1990) *FEBS Lett. 261*, 1−4.
van Duynhoven, J. P. M., Nooren, I. M. A., Swinkels, D. W., Folkers, P. J. M., Harmsen, B. J. M., Konings, R. N. H., Tesser, G. I., & Hilbers, C. W. (1993) *Eur. J. Biochem. 216*, 507−517.
Yen, T. S. B., & Webster, R. E. (1982) *Cell 29*, 337−345.
Zabin, H. B., & Terwilliger, T. C. (1991) *J. Mol. Biol. 219*, 257−275.
Zabin, H. B., Horvath, M. P., & Terwilliger, T. C. (1991) *Biochemistry 30*, 6230−6240.
Zaman, G. J. R., Schoenmakers, J. G. G., & Konings, R. N. H. (1990) *Eur. J. Biochem. 189*, 119−124.
Zaman, G., Smetsers, A., Kaan, A., Schoenmakers, J., & Konings, R. (1991) *Biochim. Biophys. Acta 1089*, 183−192.
Zaman, G. J. R., Kaan, A. M., Schoenmakers, J. G. G., & Konings, R. N. H. (1992) *J. Bacteriol. 174*, 595−600.

BI9505023