# Increasing Protein Stability by Polar Surface Residues: Domain-Wide Consequences of Interactions Within a Loop

P. R. Pokkuluri, R. Raffen, L. Dieckman, C. Boogaard, F. J. Stevens, and M. Schiffer
Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439 USA

ABSTRACT We have examined the influence of surface hydrogen bonds on the stability of proteins by studying the effects of mutations of human immunoglobulin light chain variable domain ($V_L$). In addition to the variants Y27dD, N28F, and T94H of protein κIV Len that were previously described, we characterized mutants M4L, L27cN, L27cQ, and K39T, double mutant M4L/Y27dD, and triple mutant M4L/Y27dD/T94H. The triple mutant had an enhanced thermodynamic stability of 4.2 kcal/mol. We determined the structure of the triple mutant by x-ray diffraction and correlated the changes in stability due to the mutations with changes in the three-dimensional structure. Y27dD mutant had increased stability of Len by 2.7 kcal/mol, a large value for a single mutation. Asp27d present in CDR1 formed hydrogen bonds with the side-chain and main-chain atoms within the loop. In the case of the K39T mutant, which reduces stability by 2 kcal/mol, Lys39 in addition to forming a hydrogen bond with a carbonyl oxygen of a neighboring loop may also favorably influence the surface electrostatics of the molecule. We showed that hydrogen bonds between residues in surface loops can add to the overall stability of the $V_L$ domains. The contribution to stability is further increased if the surface residue makes more than one hydrogen bond or if it forms a hydrogen bond between neighboring turns or loops separated from each other in the amino acid sequence. Based on our experiments we suggest that stabilization of proteins might be systematically accomplished by introducing additional hydrogen bonds on the surface. These substitutions are more straightforward to predict than core-packing interactions and can be selected to avoid affecting the protein's function.

## INTRODUCTION

Protein domains have been stabilized by many different methods. These include the optimization of packing density and hydrophobicity of the protein core (Matthews, 1993; Fersht and Serrano, 1993), reduction of conformational entropy in the unfolded state (e.g., methionine to leucine mutations, Gassner et al., 1996), hydrogen bonds within the core of the protein (Myers and Pace, 1996), salt bridges (Strop and Mayo, 2000), optimization of surface charge of the molecule (Perl et al., 2000; Pace, 2000), and charge-charge interaction in the unfolded state (Pace et al., 2000). For a recent review and further references see Lee and Vasmatzis (1997).

An additional method of stabilization is to introduce hydrogen bonds on the protein surface (Predki et al., 1996; Zhou et al., 1996; Takano et al., 1999). In this paper we analyze mutants that either form or break hydrogen bonds in surface loops in the immunoglobulin light-chain variable domains. Immunoglobulin light-chain variable domain ($V_L$) is a prototype β-domain that consists of two β-sheets. Most of the $V_L$ domains form dimers, and the dimerization constants and the geometry of the interface depend on the specific residues in the interface. In some of the crystal structures the two domains of the $V_L$ dimer don't have equivalent conformations because of crystal packing. Based on the amino acid sequence subgroups of light chains were identified. Furthermore, according to sequence variability, hypervariable or complementarity-determining regions (CDRs) and framework regions were defined (for review see Kabat et al., 1991). In the nomenclature of Kabat et al. (1991), extra residues in CDR1, after residue 27, are labeled 27a to 27f.

We have previously found (Raffen et al., 1999) studying a $V_L$ domain that hydrogen bonds that are on the surface, exposed to the solvent, can provide major contributions to the stabilizing of the domain. Here, we describe additional mutants designed to break or form hydrogen bonds located on the protein surface and determined their stability. New constructs included a triple mutant that combined the previously described surface loop mutations Y27dD and T94H (Raffen et al. 1999) with an internal stabilizing mutation M4L. The stabilizing effects of the three changes were additive, resulting in an increase of $\Delta G$ from 7.7 kcal/mol to 12.1 kcal/mol in the thermodynamic stability of the β-domain. The crystal structure of the triple mutant was determined, and the consequence of the mutations on the surface hydrogen bond network was ascertained. The structure of the triple mutant also served as a basis for interpreting effects of mutations that lead to the formation or breakage of hydrogen bonds in light-chain $V_L$ mutants made by us and by other investigators.

*Received for publication 10 January 2001 and in final form 19 October 2001.*

Dr. C. Boogaard was on a sabbatical leave from Ferris State University, Big Rapids, MI.

Address reprint requests to Dr. Marianne Schiffer, Biosciences Division, Argonne National Laboratory 9700 South Cass Avenue, Bldg. 202, Argonne, IL 60439. Tel.: 630-252-3883; Fax: 630-252-5517; E-mail: mschiffer@anl.gov.



0006-3495/02/01/391/08 $2.00

## MATERIALS AND METHODS

### $V_L$ mutagenesis, expression, and purification

Construction of the vectors producing the recombinant human κIV $V_L$ protein designated Len, and mutants derived from it, has been described (Wilkins Stevens et al., 1995). Site-specific mutations of these $V_L$s were produced with the MORPH site-specific plasmid DNA mutagenesis kit (5 Prime to 3 Prime, Inc., Boulder, CO), by recombinant polymerase chain reaction (Higuchi et al., 1988), or by the method of Kunkel et al. (1987). All mutations were verified by dideoxy sequencing (Sanger and Coulson, 1975). Expression and purification of $V_L$s were performed as described (Wilkins Stevens et al., 1995) except that the *Escherichia coli* host strain JM83 (Vieira and Messing, 1982) was used to improve expression of certain $V_L$s. Extinction coefficients for purified $V_L$s were calculated from the amino acid sequence with the Wisconsin Sequence Analysis Package (Version 8, September 1994; Genetics Computer Group, Madison, WI). Light-chain residues are numbered and the CDRs are defined according to Kabat et al. (1991).

### Determination of $V_L$ dimerization

$V_L$ dimerization was measured by small-zone size-exclusion HPLC using a 0.3- × 25-cm silanized glass column packed with Superdex 75 HR (Pharmacia, Piscataway, NJ). All experiments were performed at room temperature in 20 mM potassium phosphate, pH 7.0, containing 100 mM NaCl. The flow rate (0.06 ml/min) was controlled with an LKB 2150 HPLC pump; 5-μl samples were applied to the column with a Rheodyne model 7410 injector, and elution was monitored by measuring the absorbance at 214 nm with a Pharmacia UV-M monitor. Association constants for $V_L$ dimers were determined by fitting simulated elution profiles to the experimental data as described previously (Stevens and Schiffer, 1981; Stevens, 1986, 1989; Raffen and Stevens, 1999).

### $V_L$ stability measurements

Equilibrium guanidine hydrochloride (GdnHCl) denaturation of $V_L$s was measured by following the increase in tryptophan fluorescence that occurs upon exposure of Trp35, which is highly quenched in the native form. Samples containing 1.5 μM $V_L$, 10 mM sodium phosphate, pH 7.5, and various concentrations of GdnHCl were incubated overnight at 25°C. Fluorescence was measured at 25°C with an SLM Aminco SPF-500C spectrofluorimeter (SLM Instruments, Urbana, IL) with excitation at 295 nm and emission at 350 nm. Raw data were corrected for buffer fluorescence, and the denaturation curves were analyzed by the linear extrapolation method using the equation derived by Santoro and Bolen (1988). Nonlinear least-squares fitting, performed with the program KaleidaGraph (Synergy Software, Reading, PA), yielded values for $\Delta G^{o}_{unf}$, the free energy of unfolding in the absence of denaturant; $m$, the free energy change per mole of denaturant added; and the statistical error associated with these parameters. The concentration of GdnHCl at the midpoint of the denaturation curve ($C_m$) was given by $\Delta G^{o}_{unf}/m$. The change in stability, $\Delta\Delta G^{o}_{unf}$, was calculated relative to the wild-type protein by multiplying $\Delta C_m$, the difference between the mutant and wild-type $C_m$, by the $m$ value for the wild-type protein (Cupo and Pace, 1983).

### Structure determination of the triple mutant (M4L/Y27dD/T94H)

The triple mutant was crystallized by the hanging drop method from 20% PEG 3350, 10% 2-propanol, 0.1 M sodium HEPES, pH 7.5. The protein concentration was 10 mg/ml. Diffraction data to 2.7 Å were collected on an R-Axis IIc image plate detector mounted on a Rigaku RU-H2R rotating anode. Higher resolution data to 1.5 Å were collected at the Structural Biology Center (SBC), 19ID beamline of the Advanced Photon Source. Before flash cooling for data collection, crystals were briefly soaked in a cryoprotectant solution containing 30% PEG 3350 in addition to all other constituents (except the protein) of the mother liquor in which the crystals were grown. The diffraction data were processed with Denzo (Otwinowski and Minor, 1997), then merged and scaled using Scalepack (Otwinowski and Minor, 1997). Characterization of the crystals, data collection, and refinement statistics are shown in Table 1.

**TABLE 1 Crystallographic data of the M4L/Y27dD/T94H triple mutant of protein Len**

| | |
|---|---|
| Space group | $P2_12_12$ |
| a (Å) | 62.5 |
| b (Å) | 105.2 |
| c (Å) | 42.3 |
| $V_m$ (Å$^3$/dalton) | 2.8 |
| Highest resolution (Å) | 1.5 |
| Completion (%) (last shell) | 98.8 (95.6) |
| $R_{merge}$ (%) (last shell) | 4.9 (20.7) |
| Number of reflections used | 41634 |
| Final R-factor | 0.211 |
| Free R-factor (10% data) | 0.238 |
| Number of refined nonhydrogen protein atoms | 1739 |
| Number of water molecules | 361 |
| RMSD bond length (Å) | 0.005 |
| RMSD bond angle (°) | 1.3 |
| Mean temperature factor | |
| For protein atoms (Å$^2$) | 18.9 |
| For water molecules (Å$^2$) | 35.2 |

The structure was determined by molecular replacement using the program AMoRe (Navaza, 1994) with the dimer of native Len (PDB code 1LVE) as a search model. The Len dimer was generated from the monomer by rotation about the crystallographic-twofold axis. Side-chain atoms beyond CB were removed for the three mutated residues. A rotation and translation search yielded a weak solution (correlation coefficient of 23.9% and R-factor was 51.0% for 10–4-Å data). Rigid body refinement of the orientation and position of the starting dimer as a rigid body slightly improved the correlation coefficient to 28.4% and R-factor to 49.5% for 10–4-Å data. The structure was then further refined with the program CNS (Brünger et al., 1998). Refinement of the two domains of the dimer as separate rigid bodies lowered the R-factor to 36.9% (8–3-Å data). At this point, higher-resolution (1.5 Å) synchrotron data were obtained, and the model was further refined with CNS (Brünger et al., 1998). B-factor and bulk solvent corrections were applied. Manual rebuilding was performed with the program CHAIN (Sack, 1988). Resolution of the data used in refinement was increased in several small steps to 1.5 Å. After several cycles of manual adjustments and refinement, the final R-factor was 21.1% and R-free was 23.8%. The Ramachandran plot of the final model contained 90% of the residues in most favored regions with only Ala51 of both monomers in the disallowed region as observed previously (Steipe et al., 1992). The coordinates and structure factors were deposited in the Protein Data Bank, code 1EEQ.

## RESULTS

### Mutants of $V_L$ domain Len and their stabilities

Mutants Y27dD, T94H, and N28F were made and characterized previously (Raffen et al., 1999). Mutations Y27dD and T94H increased the stability ($\Delta G$) of the wild type by 2.7 kcal/mol and 0.7 kcal/mol, respectively; N28F mutation decreased the stability of the wild type. These mutations are

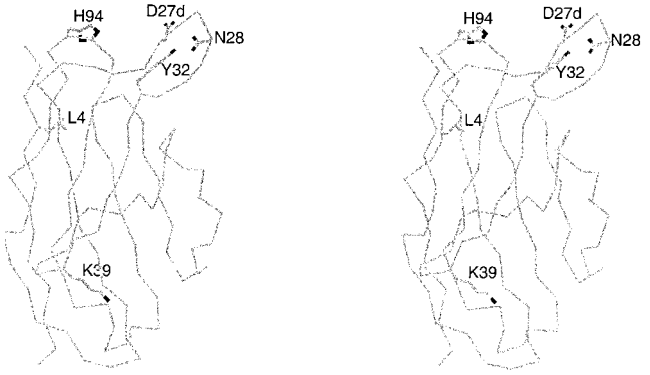


FIGURE 1 Location of the mutated residues 4, 27d, 28, 32, 39, and 94 of $V_L$ domain Len discussed in the text are shown in a stereo plot.

substitutions that affect hydrogen bonding. Additional mutants K39T, L27cN, and L27cQ on the protein surface were constructed and characterized. These mutants were constructed to further evaluate the influence of surface hydrogen bonds on protein stability. Mutation K39T, which broke a hydrogen bond to a backbone carbonyl oxygen, was destabilizing by 2 kcal/mol. The replacement of Leu on the protein surface at position 27c by potentially hydrogen-bonding residues Asn or Gln resulted in opposite outcomes. Whereas protein L27cN was destabilized by 0.4 kcal/mol, protein L27cQ was stabilized by 1 kcal/mol. An internal mutation M4L, where Met in the protein core is replaced by a Leu residue, increased stability by 1 kcal/mol.

A double mutant M4L/Y27dD and a triple mutant M4L/Y27dD/T94H were also constructed and characterized. These mutations, as shown in Fig. 1, are distant from each other. The above mutations increased the stability of the wild type by 3.5 kcal/mol and 4.2 kcal/mol respectively. These are surprisingly large stabilization values for two or three mutations; they also indicate that the contributions of the individual mutations to the stability of the Len protein are largely additive. The mutants and their stabilities are listed in Table 2. The crystal structure of the triple mutant

**TABLE 2 Stability of the Len protein mutants**

| | $\Delta G_{unf}$ (kcal/mol) | $-m$ (kcal/mol/M) | $C_m$ [GdnHCl] | $\Delta\Delta G_{unf}$ (kcal/mol)* |
|---|---|---|---|---|
| Wild type (Len)† | 7.7 ± 0.4 | 4.4 ± 0.2 | 1.76 | |
| Y27dD† | 9.4 ± 0.5 | 4.0 ± 0.2 | 2.37 | −2.7 |
| T94H† | 8.4 ± 0.7 | 4.3 ± 0.4 | 1.93 | −0.7 |
| M4L | 8.8 ± 0.6 | 4.4 ± 0.3 | 1.98 | −1.0 |
| M4L/Y27dD | 9.8 ± 0.5 | 3.8 ± 0.2 | 2.55 | −3.5 |
| M4L/Y27dD/T94H | 12.1 ± 0.2 | 4.5 ± 0.1 | 2.71 | −4.2 |
| N28F† | 4.6 ± 0.6 | 3.3 ± 0.4 | 1.37 | 1.7 |
| K39T | 5.1 ± 0.4 | 3.9 ± 0.3 | 1.31 | 2.0 |
| L27cN | 6.9 ± 0.4 | 4.1 ± 0.2 | 1.68 | 0.4 |
| L27cQ | 8.0 ± 0.5 | 4.0 ± 0.2 | 1.98 | −1.0 |

*$\Delta\Delta G_{unf} = m^{Len} \times C_m^{mut} - m^{Len} \times C_m^{Len}$.
†From Raffen *et al*. (1999).

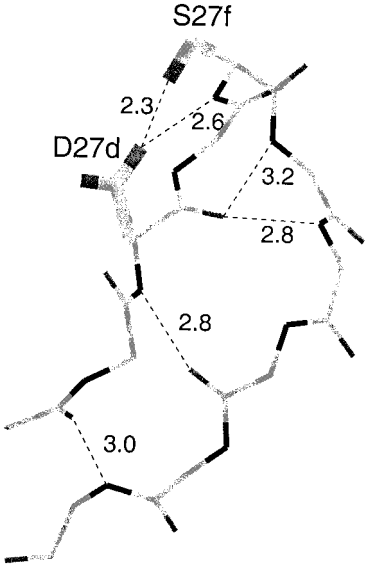


FIGURE 2 Hydrogen bonds formed by Asp27d mutant are illustrated. The backbone hydrogen bonds of the loop are also shown. The Tyr-to-Asp mutation of residue 27d increased the stability of the molecule by 2.7 kcal/mol.

was determined to see the structural influences of the introduced mutations.

## Structure of the $V_L$ domains of the triple mutant (M4L/Y27dD/T94H)

The structures of the domains of the (M4L/Y27dD/T94H) triple mutant dimer are similar to the recombinant, wild-type protein Len (rLen) (Pokkuluri et al., 1998). The root mean square (rms) deviations between $\alpha$-carbon positions of the individual domains of the triple mutant and rLen are 0.55 and 0.63 Å, respectively. The main differences are limited to two chain segments: residues 39–41 and residues 59–60. Residue 60 in protein rLen takes part in a crystal contact. The alterations of the conformations of residues 39–41 might reflect somewhat different dimer structures, although the reasons for the differences are not clear. These residues are located at the tip of the second framework loop, a loop that has been previously observed to be able to change conformation (Wang et al., 1979).

The electron densities for the three mutated residues were well defined in both monomers. Leu4 occupies the same space as Met4 in the wild-type structure. The CD2 atom of leucine is in approximately the same position as the sulfur atom of the methionine. Asp27d (OD1) forms hydrogen bonds (see Fig. 2) with the main-chain nitrogen and the side chain of Ser27f in both monomers. In monomer 1 it also forms a hydrogen bond with the peptide nitrogen of residue 28 (Table 3). Within the same monomer the ring of His94 is approximately parallel to the ring of Tyr96; the two rings

**TABLE 3 Intradomain hydrogen bonds in the triple mutant Len(M4L/Y27dD/T94H) structure discussed in the text**

| | | Monomer 1 $d$ (Å) | Monomer 2 $d$ (Å) |
|---|---|---|---|
| Asp27d OD1 | Ser27f N | 2.78 | 2.68 |
| Asp27d OD1 | Ser27f OG | 2.36 | 2.36 |
| Asp27d OD1 | Asn28 N | 3.18 | >3.50 |
| Asn28 OD1 | Lys30 N | 2.78 | 3.27 |
| Asn28 ND2 | Tyr32 OH | 3.02 | 3.09 |
| Lys39NZ | Glu81O | 3.11 | 2.83 |

overlap to some extent. Residue His94 also forms an interdomain charged hydrogen bond that will be discussed later.

Two of the mutations, N28F and K39T, reduced the stability of the Len $V_L$ domain. Residues Asn28 and Lys39 form hydrogen bonds in the wild type as well as in the triple mutant $V_L$ domains. The side chain of residue Asn28 is hydrogen bonded to Tyr32, effectively linking the tip of CDR1 (Asn28) to its base (Tyr32) (Fig. 3). Asn 28 also forms a hydrogen bond with the peptide nitrogen of residue 30. The side chain of Lys39 forms a hydrogen bond with the backbone carbonyl of residue 81, cross-linking two loops of the domain.

## Domain interactions within the M4L/Y27dD/T94H dimer

The triple mutant forms a non-crystallographic $V_L$ domain dimer in the crystal. The angle of rotation around the non-crystallographic twofold axis is 179.1°. The $V_L$ domains have very similar structures; the rms deviation between $\alpha$-carbons upon overlapping the monomers is 0.29 Å. The largest difference occurs for serine residues 76 and 77; these residues are not well defined, and they have large temperature factors in both domains. The dimers formed by the triple mutant and the wild-type protein differ, when one of the domains of the triple mutant and the wild-type structure are superimposed, an 11° rotation and a translation of 2.5 Å are required to superimpose the second domains. The differences in the dimer structures are caused by the introduction of extra interdomain hydrogen bonds in the triple mutant. As a result of the extra hydrogen bonds, the dimerization constant of $2 \times 10^5$ $M^{-1}$ of the wild type increased 400-fold to $8 \times 10^7$ $M^{-1}$ in the triple mutant.

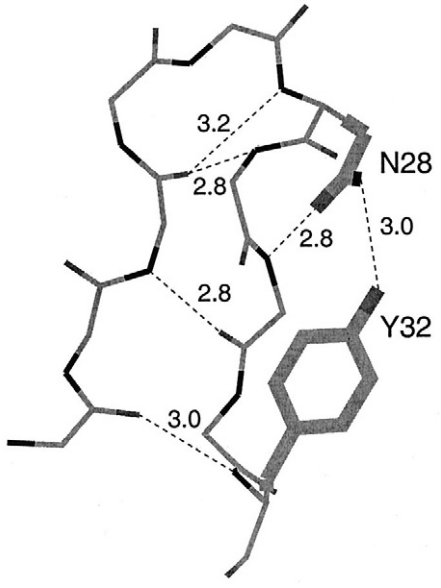


FIGURE 3 The hydrogen bond between Asn28 and Tyr32 is shown. The side chain of Asn28 also forms a hydrogen bond with the backbone nitrogen of residue 30. The Asn-to-Phe mutation of residue 28 removes the above two hydrogen bonds and introduces a hydrophobic residue on the surface; it reduces the stability of the molecule by 1.7 kcal/mol.

**TABLE 4 Interdomain hydrogen bonds**

| Monomer 1 | Monomer 2* | Triple mutant $d$ (Å) | rLen† $d$ (Å) |
|---|---|---|---|
| Gln38 OE1 | NE2 338 Gln | 2.83 | 2.82 |
| Gln38 NE2 | OE1 338 Gln | 2.85 | 2.82 |
| Glu55 OE1 | ND1 394 His | 2.86 | |
| His94 ND1 | OE1 355 Glu | 2.72 | |
| Glu55 OE2 | N 396 Tyr | 2.74 | 5.92 |
| Tyr96 N | OE1 355 Glu | 3.19 | 5.22 |
| Tyr36 OH | NE2 389 Gln | 3.30 | 5.32 |
| Gln89 NE2 | OH 336 Tyr | 3.19 | 5.32 |
| Gln42 O | OH 387 Tyr | 3.45 | 3.48 |
| Tyr87 OH | O 342 Gln | 3.48 | 3.48 |
| Gln38 NE2 | OH 387 Tyr | 3.58 | 3.20 |
| Tyr87 OH | NE2 338 Gln | 3.63 | 3.20 |

Hydrogen bonds with $d < 3.5$ Å in one of the structures are listed.
*Three hundred was added to the numbers of the second monomer.
†The recombinant Len protein occurs as a crystallographic dimer. Residue 94 is a threonine.

A comparison of interdomain hydrogen bonds between the triple mutant and the wild-type Len dimers is shown in Table 4 (300 was added to the residue numbers of the second monomer). The interdomain hydrogen bonds formed by residues Gln38 are the same in both. The newly introduced His94 of the mutant forms a charged hydrogen bond with the side chain of Glu55 (Glu355) of the other monomer. Glu355 also forms an additional interdomain hydrogen bond to the backbone nitrogen of residue 96. The hydrogen bonding by residues Glu55 and Glu355 are not the same. The OE1 atom of Glu55 forms a hydrogen bond with His394, and the OE2 atom of Glu55 forms hydrogen bond with the peptide nitrogen of residue 396. In contrast, both hydrogen bonds by Glu355 are formed by the same atom, OE2. For comparison, in the rLen structure, atoms OE1 and OE2 of Glu55 are 5.22 and 5.92 Å from the backbone nitrogen atom of residue 96, too distant to form hydrogen bonds. Therefore, to achieve the interdomain hydrogen bonds between residues Glu55 and His94 on both sides of the dimer in the triple mutant, the orientations of the domains in the triple mutant had to change relative to the wild

type. The distance between $\alpha$-carbons of 55 and 394 and 355 and 94 are 1.4 Å longer in rLen than in the triple mutant, whereas the interdomain distances between $\alpha$-carbons of 38 and 387, and 338 and 87, are longer in the triple mutant than in the rLen. The distance between the $\alpha$-carbons of residue 36 and 336 in the two structures remains the same.

## DISCUSSION

### Evaluation of the stabilizing mutations of the triple mutant (M4L/Y27dD/T94H) based on its three-dimensional structure

We have made mutations in the immunoglobulin $V_L$ domain that changes hydrogen bonding by residues at or near the surface and made a Met-to-Leu substitution in the core of the domain. We determined the effect of the mutations on the stability of the domains and determined and examined the structure of a stabilized $V_L$ domain to understand the influence of the mutations. Three mutations were introduced that were found to be stabilizing as single-site mutants. As shown in Table 2 the M4L, Y27dD, and T94H mutations increased the stability of the wild type by 1.0, 2.7, and 0.7 kcal/mol, respectively. Introduction of two mutations (M4L/Y27dD) increased the stability by 3.5 kcal/mol. The introduction of all three mutations (M4L/Y27dD/T94H) increases stability by 4.2 kcal/mol to 12.1 kcal/mol. The mutations M4L, Y27dD, and T94H are distant from each other in the three-dimensional structure; they can be combined, and their effect on the stability is additive as has been shown for other systems (e.g., Wells, 1990).

Two of the mutated residues 27dD and 94H are on the surface of the domain. They are located in the CDR1 and CDR3 loops, respectively. Mutations Y27dD and T94H, though observed in amyloidogenic light chains, increase stability by 2.7 kcal/mol and 0.7 kcal/mol, respectively, relative to wild-type protein Len (Raffen et al., 1999). The unexpected large increase of 2.7 kcal/mol of stability by the Y27dD mutation is the result of increased numbers of hydrogen bonds formed by the side chain of residue Asp27d, compared with that of the Tyr residue at position 27d. Asp27d forms several hydrogen bonds with both the backbone and side chains of residues within the same loop (see Fig. 2 and Table 3). In addition to hydrogen bond formation, this mutation also introduces a negative charge that also could affect stability.

The Y27dD mutation clearly shows that stabilization of a surface loop stabilizes the entire domain. The T94H mutation appears to stabilize the protein by both stabilizing the individual domains and the dimer they form. The individual domains might be stabilized by the interactions between aromatic rings (Fernández-Recio et al., 1999) of the His94 and Tyr96 residues. The dimerization ability of the triple mutant was enhanced by a factor of 400 from $2 \times 10^5\ M^{-1}$ to $8 \times 10^7\ M^{-1}$. At present we cannot separate the influence of T94H mutation on domain stability from its effect on the enhancement of dimerization. The increase in dimerization is the result of the two hydrogen bonds formed by residue Glu55 from one monomer with the mutated residue His94 and the backbone nitrogen of residue 96 of the second monomer of the dimer. These results can be compared with data from a different human $V_L$ dimer, Rei. In that molecule, histidine at position 94 in the L94H mutant increased the dimerization constant 20-fold from $2.5 \times 10^5\ M^{-1}$ to $5 \times 10^6\ M^{-1}$ and did not influence the stability of the protein (Kolmar et al., 1994). The change in amino acid sequence that accounts for the different influence of His94 on the dimerization of the two $V_L$ dimers is that in $\kappa$I Rei protein, the residue 55 is a neutral glutamine instead of glutamic acid, which cannot form the two charged hydrogen bonds observed in the Len triple mutant. The structure of the L94H mutant of protein Rei is not known; therefore the orientations of the His94 relative to Tyr96 residues within the same monomer cannot be compared.

Residue Met4 is part of the hydrophobic core. The M4L mutation was identified previously as stabilizing by Benhar and Pastan (1995) in a murine Fv fragment. We found that the Met-to-Leu mutation increases stability of the $V_L$ by 1 kcal/mol. The leucine residue is completely buried, and it occupies the same space as the methionine in the wild-type protein. In the interior of a protein a Met-to-Leu substitution is expected to stabilize the protein by $\sim$1.4 kcal/mol if it does not introduce strain in the structure (Gassner et al., 1996). This increase in stability was not achieved in T4 lysozyme mutants, where the stability increase was only 0.5 kcal/mol (Lipscomb et al., 1998).

### Energetic effects of additional surface hydrogen bonds of the $V_L$ domain

The breaking of hydrogen bonds at the domain surface diminished the stability of the $V_L$ domain. We determined the loss of stability when the hydrogen bonds between the side chains of Asn28 and Tyr32 (Fig. 3) and the side-chain Lys39 and the carbonyl oxygen of residue 81 were broken. The N28F and the K39T mutants diminished the domain stability by 1.7 and 2.0 kcal/mol, respectively (Table 2). The energetic contributions of these hydrogen bonds have also been reported in the murine light chain McPC603 and human $\kappa$I light chain Rei, respectively (Steipe et al., 1994; Frisch et al., 1996; Usón et al., 1997, 1999).

The mutation of Asn28 to Phe has several consequences (Fig. 4). The N28F mutation 1) breaks the Asn28 to Tyr32 hydrogen bond, 2) breaks the hydrogen bond between Asn28 and peptide nitrogen of residue 30, and 3) introduces Phe at a solvent-exposed location. The N28F mutant is 1.7 kcal/mol less stable than the wild type in the $\kappa$IV Len background (Raffen et al., 1999). The creation of a similar hydrogen bond between Asn28 and Tyr32 was stabilizing in

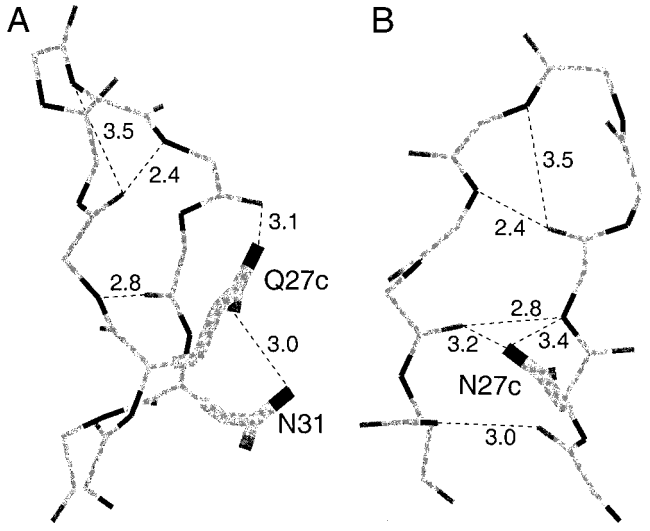


FIGURE 4 Modeled structures of L27cQ and L27cN mutants. (*A*) The two possible hydrogen bonds formed by Gln27c with the backbone carbonyl oxygen of residue 29 and with the side chain of Asn31. This mutant increases stability of the Len $V_L$ domain by 1.0 kcal/mol. (*B*) The possible hydrogen bond formed by Asn27c, but this conformation of the residue brings it too close to the peptide nitrogen of residue 27d. L27cN mutant is 0.4 kcal/mol less stable than its parent.

the murine light chain McPC603 but less so (Steipe et al., 1994). The hydrogen bond in McPC603 was created by mutation of Phe32 to Tyr; it stabilized the domain by 0.4 kcal/mol. The difference in the apparent contribution to stability of the same, Asn28 to Tyr32, hydrogen bond in proteins Len and McPC603, is caused by mutating different members of the hydrogen-bonding pairs. The N28F mutation of protein Len broke two hydrogen bonds, one to Tyr32 and one to the main chain, and introduced a hydrophobic residue on the surface, whereas the F32Y mutation of McPC603 removed a hydrophobic residue from the surface but created only one extra hydrogen bond.

In the κIV Len background the K39T mutation, which breaks the hydrogen bond formed by Lys residue 39 with the main-chain carbonyl of residue 81 located on a neighboring loop, results in a protein that is 2 kcal/mol less stable than the wild type. A similar observation was made on the influence of this hydrogen bond on stability in the κI light chain Rei where the creation of the hydrogen bond by the T39K mutation increased the stability of the $V_L$ domain by 1.3 kcal/mol (Frisch et al., 1996; Usón et al., 1999). In the Rei structure (Epp et al., 1975), Thr39 is not directly hydrogen bonded to any protein atoms but forms a hydrogen bond with a water molecule that interacts with main-chain oxygen atoms of Val83 and Glu81. In addition to forming a hydrogen bond, Lys39 is also close to two acidic residues, Glu81 and Asp82. Therefore, the stabilization effect can have both hydrogen-bonding and charge components. The stabilization of the Thr-to-Lys mutation in protein Rei is less compared with Len probably because the Rei protein surface is more positive near the mutation site. It is further suggestive that a positively charged residue is required at position 39 of the $V_L$ domain, because most of the residues at position 39 are either Lys or Arg in the immunoglobulin database (Kabat et al., 1991).

In the Rei $V_L$ domain an additional stabilizing hydrogen bond was also observed. The Tyr32-to-His mutation increased the stability of the Rei $V_L$ by 1.3 kcal/mol (Usón et al., 1997). His32 forms a hydrogen bond with Glu50; the two residues are located on neighboring loops in the structure.

## Design of a stabilizing surface hydrogen bond for the $V_L$ domain

To this point, we and others have shown that surface hydrogen bonds can stabilize $V_L$ domains, by introduction of hydrogen-bonding residues that were previously observed in other $V_L$ sequences. We identified de novo mutations that would lead to additional hydrogen bonds that might stabilize the structure. Based on computer modeling, residue Leu27c, a hydrophobic residue on the surface, was replaced by polar amino acids Asn and Gln (Fig. 4). The effects of these Leu27c mutations were expected to be twofold: they 1) would eliminate a solvent-exposed hydrophobic residue and 2) make hydrogen bond formation possible. Asn was modeled to interact with carbonyl oxygen of residue 30; Gln was modeled to interact with both the backbone carbonyl oxygen of 29 and the side chain of Asn31. Indeed, an increase of 1.0 kcal/mol stability was observed for the Gln27c mutant, but a decrease of 0.4 kcal/mol in stability was observed for the Asn27c mutant. The Gln27c mutant had the characteristics we expected, but the Asn27c is actually destabilizing; it might interact unfavorably with the protein backbone. Further examination of the modeled structure indicated that the decrease of stability of the Asn27c mutant is probably caused by a conflict of atom ND2 of Asn27c with the backbone nitrogen of residue 27d in the orientation where it can form a hydrogen bond (Fig. 4).

## Effect of surface hydrogen bonds on stability

Clearly, our experiments and data from other studies show that hydrogen bonds between residues in surface loops can add to the overall stability of the $V_L$ domain. Because these hydrogen-bonding residues are in contact with the solvent, it is expected that the hydrogen bonds are present only part of the time, as the residues can form transient interactions with the solvent. When the mutated residue forms or breaks more than one hydrogen bond with protein atoms as we show for the Y27dD and N28F, there is an increased energy gain or loss. This may be rationalized on the basis of cooperativity between multiple intramolecular hydrogen

bonds, limiting the chance for hydrogen bond formation with the solvent. Therefore, stabilizing hydrogen bonds are expected to be present a larger fraction of the time. The stabilization energy gain is also increased if the hydrogen bond is formed between a side chain and main chain, especially if the hydrogen bond stabilizes neighboring turns or loops that are separated from each other in the amino acid sequence.

## Biological implications

Interestingly, protein loops are often involved in protein activities, including ligand binding and catalysis. As such, certain loop sequences and their interactions might represent the results of evolutionary compromise between structural stability and functional robustness. Only two of the seven variants described here involve residues of the immunoglobulin framework regions, the portion of $V_L$ conventionally associated with the protein's fold and stability. Five of the replacements are located in CDRs, those portions of the molecule that undergo extensive somatic mutation during the natural course of generating immune system diversity. Thus, wide variation in stabilities of antibodies may be expected. This is borne out in the often low stabilities of monoclonal antibodies and scFvs (Benhar and Pastan, 1995; Jung and Plückthun, 1997) as well as in physiological aggregation of light chains to form amyloid deposits, which has fatal consequences (Bellotti et al., 2000; Stevens, 2000). The extensive primary structure variation of the immunoglobulin variable domains coupled with conservation of its three-dimensional structure suggests that these proteins provide a useful system for detailed study of structure/stability relationships.

This work was supported by the U.S. Department of Energy, the Office of Biological and Environmental Research under contract W-31–109-Eng-38, and by the U.S. Public Health Service grant DK43757. Use of the Argonne National Laboratory Structural Biology Center beam line at the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Energy Research, under contract W-31–109-Eng-38. We appreciate the help of SBC staff in data collection.

## REFERENCES

Bellotti, V., P. Mangione, and G. Merlini. 2000. Immunoglobulin light chain amyloidosis: the archetype of structural and pathogenic variability. *J. Struct. Biol.* 130:280–289.

Benhar, I., and I. Pastan. 1995. Identification of residues that stabilize the single-chain Fv of monoclonal antibodies B3. *J. Biol. Chem.* 270: 23373–23380.

Brünger, A. T., P. D. Adams, G. M. Clore, W. L. Delano, P. Gros, R. W. Grosse-Kunstleve, J-S. Jiang, J. Kuszewski, M. Nigles, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren. 1998. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta. Cryst.* D54:905–921.

Cupo, J. F., and C. N. Pace. 1983. Conformational stability of mixed disulfide derivatives of β-lactoglobulin B. *Biochemistry.* 22:2654–2658.

Epp, O., E. L. Lattman, M. Schiffer, R. Huber, and W. Palm. 1975. The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein Rei refined at 2.0-Å resolution. *Biochemistry.* 14: 4943–4952.

Fernández-Recio, J., A. Romero, and J. Sancho. 1999. Energetics of a hydrogen bond (charged and neutral) and of a cation-π interaction in apoflavodoxin. *J. Mol. Biol.* 290:319–330.

Fersht, A. R., and L. Serrano. 1993. Principles of protein stability derived from protein engineering experiments. *Curr. Opin. Struct. Biol.* 3:75–83.

Frisch, C., H. Kolmar, A. Schmidt, G. Kleemann, A. Reinhardt, E. Pohl, I. Usón, T. R. Schneider, and H.-J. Fritz. 1996. Contribution of the intramolecular disulfide bridge to the folding stability or $REI_V$, the variable domain of a human immunoglobulin κ light chain. *Folding Design.* 1:431–440.

Gassner, N. C., W. A. Baase, and B. W. Matthews. 1996. A test of the "jigsaw puzzle" model for protein folding by multiple methionine substitutions within the core of T4 lysozyme. *Proc. Natl. Acad. Sci. U.S.A.* 93:12155–12158.

Higuchi, R., B. Krummel, and R. K. Saiki. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* 16:7351–7367.

Jung, S., and A. Plückthun. 1997. Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng.* 10:959–966.

Kabat, E. A., T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller. 1991. Sequences of Proteins of Immunological Interest, 5th ed. NIH publication 91–3242. U.S. Department of Health and Human Services, U.S. Government Printing Office, Washington, DC.

Kolmar, H., C. Frisch, G. Kleemann, K. Götze, F. J. Stevens, and H.-J. Fritz. 1994. Dimerization of Bence Jones proteins: linking the rate of transcription from an *Escherichia coli* promoter to the association constant of $REI_V$. *Biol. Chem. Hoppe-Seyler.* 375:61–70.

Kunkel, T. A., J. D. Roberts, and R. A. Zakour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154:367–382.

Lee, B., and G. Vasmatzis. 1997. Stabilization of protein structures. *Curr. Opin. Biotechnol.* 8:423–428.

Lipscomb, L. A., N. C. Gassner, S. D. Snow, A. M. Eldridge, W. A. Baase, D. L. Drew, and B. W. Matthews.1998. Context-dependent protein stabilization by methionine-to-leucine substitution shown in T4 lysozyme. *Protein Sci.* 7:765–773.

Matthews, B. W. 1993. Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* 62:139–160.

Myers, J. K., and C. N. Pace. 1996. Hydrogen bonding stabilizes globular proteins. *Biophys. J.* 71:2033–2039.

Navaza, J. 1994. AMoRe: an automated package for molecular replacement. *Acta Cryst.* A50:157–163.

Otwinowski, Z., and W. Minor. 1997. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276:307–326.

Pace, C. N. 2000. Single surface stabilizer. *Nat. Struct. Biol.* 7:345–346.

Perl, D., U. Mueller, U. Heinemann, and F. X. Schmid. 2000. Two exposed amino acid residues confer thermostability on a cold shock protein. *Nat. Struct. Biol.* 7:380–383.

Pokkuluri, P. R., D-B. Huang, R. Raffen, X. Cai, G. Johnson, P. Wilkins Stevens, F. J. Stevens, and M. Schiffer. 1998. A domain flip as a result of a single amino-acid substitution. *Structure.* 6:1067–1073.

Predki, P. F., V. Agrawal, A. T. Brünger, and L. Regan. 1996. Amino-acid substitutions in a surface turn modulate protein stability. *Nat. Struct. Biol.* 3:54–58.

Raffen, R., L. J. Dieckman, M. Szpunar, C. Wunschl, P. R. Pokkuluri, P. Dave, P. Wilkins Stevens, X. Cai, M. Schiffer, and F. J. Stevens. 1999. Physicochemical consequences of amino acid variations that contribute to fibril formation by immunoglobulin light chains. *Protein Sci.* 8:509–517.

Raffen, R., and F. J. Stevens. 1999. Small zone, high-speed gel filtration chromatography to detect protein aggregation associated with light chain pathologies. *Methods Enzymol.* 309:318–332.

Sack, J. S. 1988. CHAIN: a crystallographic modeling program. *J. Mol. Graph.* 6:224–225.

Sanger, F., and A. R. Coulson. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* 94:441–448.

Santoro, M. M., and D. W. Bolen. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry.* 27:8063–8074.

Steipe, B., A. Plückthun, and R. Huber. 1992. Refined crystal structure of a recombinant immunoglobulin domain and a complementarity-determining region 1-grafted mutant. *J. Mol. Biol.* 225:739–753.

Steipe, B., B. Schiller, A. Plückthun, and S. Steinbacher. 1994. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* 240:188–192.

Stevens, F. J. 1986. Analysis of protein-protein interaction by simulation of small-zone size-exclusion chromatography: application to an antibody-antigen association. *Biochemistry.* 25:981–993.

Stevens, F. J. 1989. Analysis of protein-protein interaction by simulation of small-zone size exclusion chromatography. *Biophys. J.* 55:1155–1167.

Stevens, F. J. 2000. Four structural risk factors identify most fibril-forming kappa light chains. *Amyloid Int. J. Exp. Clin. Invest.* 7:200–211.

Stevens, F. J., and M. Schiffer. 1981. Computer simulation of protein self-association during small-zone gel filtration: estimation of equilibrium constant. *Biochem. J.* 195:213–219.

Strop, P., and S. L. Mayo. 2000. Contribution of surface salt bridges to protein stability. *Biochemistry.* 39:1251–1255.

Takano, K. Y. Yamagata, M. Kubota, J. Funahashi, S. Fujii, and K. Yutani. 1999. Contribution of hydrogen bonds to the conformational stability of human lysozyme: calorimetry and x-ray analysis of six Ser→Ala mutants. *Biochemistry.* 38:6623–6629.

Usón, I., M. T. Bes, G. M. Sheldrick, T. R. Schneider, T. Hartsch, and H-J. Fritz. 1997. X-ray crystallography reveals stringent conservation of protein fold after removal of the only disulfide bridge from a stabilized immunoglobulin variable domain. *Folding Design.* 2:357–361.

Usón, I., E. Pohl, T. R. Schneider, Z. Dauter, A. Schmidt, H-J. Fritz, and G. M. Sheldrick. 1999. 1.7 Å structure of the stabilized $REI_v$ mutant T39K: application of local NCS restraints. *Acta. Cryst.* D55:1158–1167.

Vieira, J., and J. Messing. 1982. The Puc plasmids, an M13Mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene.* 19:259–268.

Wang, B-C., C. S. Yoo, and M. Sax. 1979. Crystal structure of Bence Jones protein Rhe (3 Å) and its unique domain-domain association. *J. Mol. Biol.* 129:657–674.

Wells, J. A. 1990. Additivity of mutational effects in proteins. *Biochemistry.* 29:8509–8517.

Wilkins Stevens, P., R. Raffen, D. K. Hanson, Y-L. Deng, M. Berriios-Hammond, F. W. Westholm, C. Murphy, M. Eulitz, R. Wetzel, A. Solomon, M. Schiffer, and F. J. Stevens. 1995. Recombinant immunoglobulin variable domains generated from synthetic genes provide a system for in vitro characterization of light chain amyloid proteins. *Protein Sci.* 4:421–432.

Zhou, H. X., R. H. Hoess, and W. F. DeGrado. 1996. In vitro evolution of thermodynamically stable turns. *Nat. Struct. Biol.* 3:446–451.