Mutagenesis of a Buried Polar Interaction in an SH3 Domain: Sequence Conservation Provides the Best Prediction of Stability Effects†

Karen L. Maxwell‡ and Alan R. Davidson*,‡,§

Department of Molecular and Medical Genetics and Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A8

Received July 24, 1998; Revised Manuscript Received September 14, 1998

ABSTRACT: The SH3 domain from the Fyn tyrosine kinase possesses a buried hydrogen bond between the side chains of a glutamate (Glu24) and a serine (Ser41) residue. Multiple amino acid substitutions were made at these positions to determine the role of this interaction in the stability and conformational specificity of the domain and to assess the relationship between the thermodynamic stability of mutants and sequence conservation seen in the SH3 domain family. Analysis of single and double alanine mutations indicated that the Glu24−Ser41 interaction contributes 0.50 kcal/mol to the stability of the domain. However, disruption of the Glu24−Ser41 interaction did not impair peptide binding function, suggesting that the interaction is not critical for conformational specificity. The stability of the domain was not increased by the replacement of these residues with different combinations of hydrophobic residues or with potential salt bridge forming residues. Despite their similar structural roles in the Fyn SH3 domain, the Ser41 position was considerably more tolerant to substitution than was the Glu24 position. An alignment of >350 different SH3 domains has been completed in our laboratory. A statistically significant correlation was found between the conservation data for the Glu24 and Ser41 positions and the thermodynamic stabilities of the mutants constructed at these positions. Surprisingly, our analysis of sequence alignment data provided a more accurate prediction of the stability of mutants than did examination of the three-dimensional structure of the domain.

SH3 domains mediate protein−protein interactions in a wide variety of signaling pathways in eukaryotic cells (*1*, *2*). Their important biological role and amenability to biophysical techniques have made them an attractive subject for both structural and protein folding studies. Moreover, the presence of >350 different SH3 domain sequences within the protein sequence database provides a rich source of information for the design and interpretation of experiments. Determination of more than a dozen different SH3 domain structures (e.g., refs *3*−*7*) has shown that each displays essentially the same fold, consisting of two three-stranded β-sheets packed orthogonally against one another to form a single hydrophobic core (Figure 1A). SH3 domains exhibit two-state behavior when unfolded by either heat or chemical denaturants (*8*−*11*), allowing thermodynamic analysis of folding and quantitation of the effects of amino acid substitutions on stability. A patch of conserved aromatic residues clustered on the surface of the molecule mediates binding to proline-rich target peptides (*12*) (Figure 1A), which can be assayed by monitoring the increase in fluorescence that occurs upon binding.

In this study, we assessed the changes in thermodynamic stability and peptide-binding activity caused by amino acid substitutions at the Glu24 and Ser41 positions in the Fyn tyrosine kinase SH3 domain. In the crystal structure (*5*), the O_δ atom of Glu24 and the hydroxyl oxygen of Ser41 are only 2.9 Å apart and lie in close to ideal geometry to form a hydrogen bond (*13*) (Figure 1C,D). On the basis of their proximity (Figure 1D), hydrogen bonds are also likely formed between the other O_δ atom of Glu24 and the amide hydrogen of His21 and between the O_γ atom of Ser41 and the amide hydrogen of Thr44. Glu24 lies at the i+3 position of a type II β-turn, whereas Ser41 is positioned in the third β-strand of the domain. Glu24 and Ser41 are 86 and 99% buried, respectively, and both are buried in a 50% nonpolar environment, as calculated using ENVIRONMENTS (*14*). The nonpolar burial of these residues is primarily due to packing against Leu18, Phe20, and Phe26 in the hydrophobic core of the domain. Our mutagenic study of these positions had two objectives. The first was to assess the contribution of the buried polar interaction between Glu24 and Ser41 to overall thermodynamic stability and to determine whether this interaction is important for imparting conformational specificity to the domain. The second was to compare the effects caused by mutating Glu24 and Ser41 in the Fyn SH3 domain with the amino acid conservation observed in a large alignment of diverse SH3 domain sequences.

The role of buried polar interactions in stabilizing proteins and in conferring conformational specificity is a topic of great current interest. Virtually every known protein structure contains buried charged or polar residues, the majority of which are involved in hydrogen bonds (*15*, *16*). The

† This work was supported by Medical Research Council of Canada Grant MT-13609. This work is dedicated to Vernon Maxwell.

* To whom correspondence should be addressed (e-mail davidson@hkl.med.utoronto.ca).

‡ Department of Molecular and Medical Genetics.

§ Department of Biochemistry.

10.1021/bi981788p CCC: $15.00
Published on Web 10/28/1998

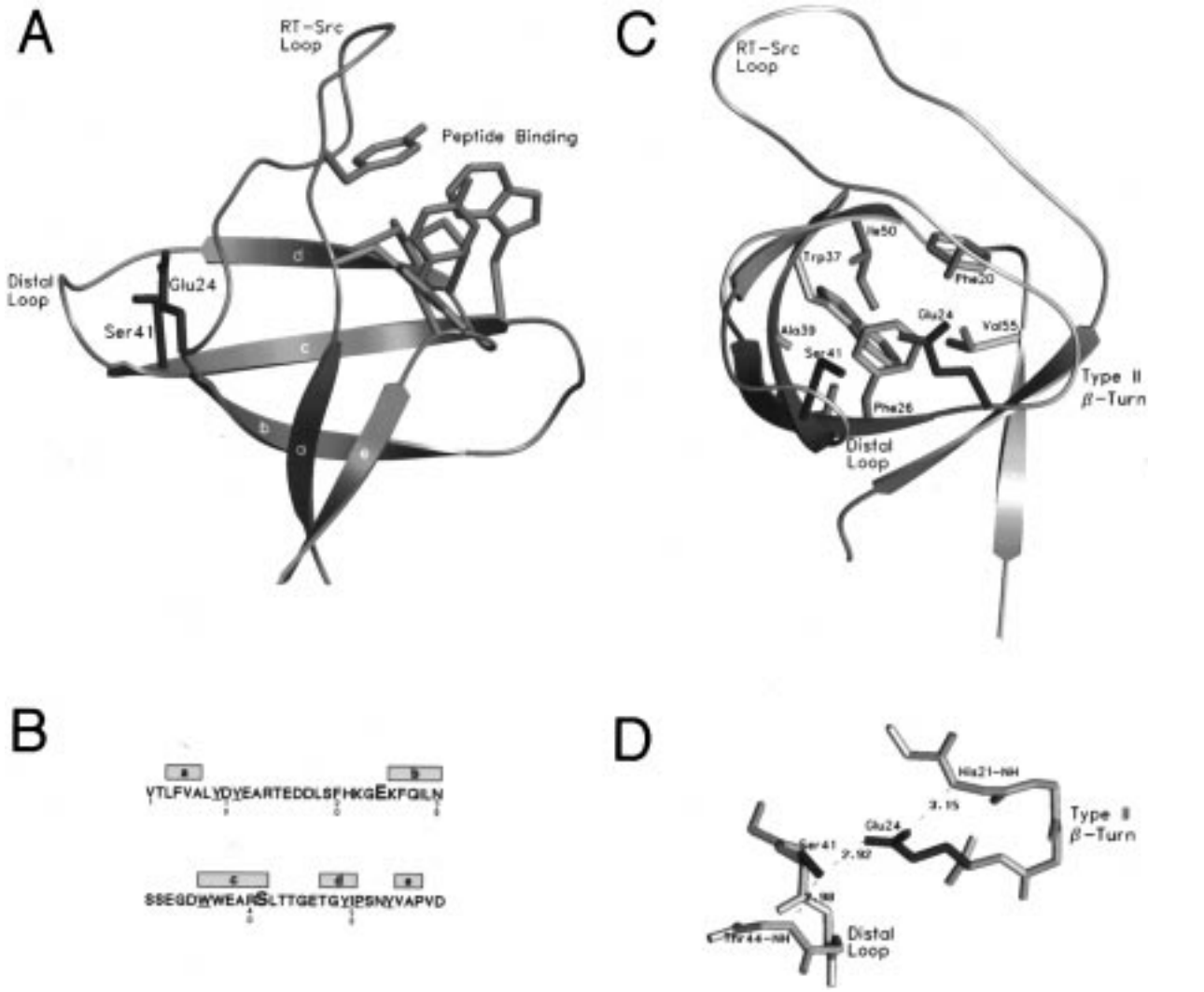

FIGURE 1: (A) Ribbon drawing of the WT Fyn SH3 domain based on the X-ray crystal structure (PDB identifier 1SHF) (*5*). The strands comprising the two β-sheets are labeled a−e. The side chains of conserved residues involved in peptide binding are shown, as are those of Glu24 and Ser41. Graphics were produced using SETOR (*49*). (B) Amino acid sequence (residues 84−142 of the Fyn tyrosine kinase) of the domain shown above. The positions of the five β-strands are indicated by boxes, and conserved residues involved in peptide binding are underlined. Glu24 and Ser41 are indicated by larger letters. (C) View of the Fyn SH3 domain showing the position of Glu24 and Ser41 with respect to some hydrophobic core residues. Compared to (A), the molecule here is rotated 90° around the *y*-axis and 45° around the *x*-axis. (D) Hydrogen bond interactions of Glu24 and Ser41. The numbers indicate the distances in angstroms between the joined atoms. The molecule is viewed at an angle similar to that above.

importance of these interactions in protein stability has been implied by a number of pK_a shift and mutagenesis experiments demonstrating that disruption of a buried polar interaction can destabilize a protein by anywhere from 1 to 5 kcal/mol (*17*−*20*). However, a continuum electrostatic study examining buried salt bridges from a variety of proteins concluded that the majority of these interactions were actually destabilizing when compared with substitution of hydrophobic isosteres at the same positions (*21*). Experimental support for this conclusion was provided by Waldburger et al., who demonstrated that substitution of a buried salt bridge triad in bacteriophage P22 arc repressor with various combinations of hydrophobic residues led to dramatic stabilization of the protein (*20*). Even if buried polar interactions are not crucial in protein stability, they may still play an important role in maintaining the unique native structure of a protein by limiting the number of allowable conformations to those in which compensating interactions can take place between the polar residues (*21*). A number of protein design experiments have supported this notion (*22*, *23*), although the replacement of the buried salt bridge in the arc repressor had no effect on conformational specificity. The Glu24−Ser41 interaction in the Fyn SH3 domain provides an excellent system to assess the role of a buried polar interaction. This interaction is one of only two potential inter-side-chain buried polar interactions in the domain (the other is between Tyr 10 and Asp 17); thus, its elimination should significantly perturb the stability of the domain if such interactions play an important role in its stability.

The Glu24 and Ser41 positions also provide an interesting test case for assessing the relationship between amino acid conservation in the SH3 domain family and the effects of mutations in an individual domain. Because these residues are clearly interacting with each other and are both highly buried in the Fyn SH3 domain structure, substitutions at the Glu24 and Ser41 positions might be expected to have similar effects and be generally destabilizing. However, our analysis of a large alignment of SH3 domain sequences has shown that the conservation patterns at these two positions are very

different. Whereas the Glu24 position is the seventh most conserved position in the family and is either Glu or Asp 85% of the time, the Ser41 position is one of the least conserved in the family. Thus, from sequence conservation data, one would predict that the Ser41 position would be much more tolerant to amino acid substitutions than the Glu24 position.

EXPERIMENTAL PROCEDURES

Plasmid Construction and Mutagenesis. cDNA encoding the chicken Fyn tyrosine kinase SH3 domain was obtained from the laboratory of David Baltimore. The SH3 domain (residues 85−142) was amplified by PCR and cloned into the expression vector pAD100 (*24*), which contains a P_{tac}-based, IPTG-inducible promoter. The SH3 domain encoding DNA was cloned in such a way that 5 additional residues (MVQIS) were added to the N terminus of the expressed protein and 16 (RLDYKDDDDKHHHHHH) were added to the C terminus. Encoded within the last 16 residues are the FLAG epitope and the hexahistidine sequence used for protein purification. Site-directed mutations were introduced into the SH3 gene by two-step site-directed PCR mutagenesis (*25*) using Vent DNA polymerase (New England Biolabs). The DNA sequences of the wild-type gene and all mutants used in this study were verified using the Sequenase version 2.0 kit (United States Biochemical). Highly destabilized mutants, which did not accumulate at high levels when expressed from the pAD100-based construct, were subcloned into pET21d (Novagen) in such a way that the SH3 domain was expressed with the same N- and C-terminal extensions as described above.

Protein Expression and Purification. All protein expression was performed in the *Escherichia coli* strain BB100, the genotype of which is *ara* Δ(*lac proAB*) Δ*slyD* (*kan*^{r}) *nalA argEam rif thi* F′[*lacI*^{q} *proAB*$^+$]. The *slyD* deletion prevents the expression of a 21 kDa histidine-rich *E. coli* protein that binds strongly to the Ni-affinity resin used to purify the proteins studied in this work. pET21d-based vectors were expressed in a derivative of BB100 lysogenized with λDE3, which contains the gene for T7 RNA polymerase under the control of the IPTG-inducible *lacUV5* promoter (*26*). After IPTG induction and 2 h of growth at 37 °C, cells were harvested and lysed in 6 M GuHCl,[1] 100 mM NaH_2PO_4, 10 mM Tris-HCl, and 10 mM imidazole, pH 8.0. Purification was carried out by a batch procedure in the same buffer using NiNTA-agarose resin (Qiagen). Purified proteins were eluted with low pH (6 M GuHCl, 0.2 M acetic acid) and were refolded by dialysis into 10 mM Tris-HCl, pH 8.0, 0.2 mM EDTA, and 250 mM KCl. All subsequent folding and functional assays were performed in this buffer. Protein purity was verified to be >98% by Coomassie staining of samples following SDS−PAGE gel electrophoresis. Protein concentrations were determined by UV absorbance at 280 nm using a molar absorption coefficient of 18 450 M^{-1} cm^{-1} (*27*).

Determination of Thermodynamic Parameters by Circular Dichroism (CD) Spectroscopy. CD measurements were performed in an Aviv 62A DS circular dichroism spectrometer. Thermal and chemical denaturation were monitored by following the change in ellipticity at 220 nm. Proteins were heated from 25 to 109 °C in 2 °C increments, with a 1 min equilibration time and a 15 s averaging time for the CD measurement. Less stable proteins (T_m < 60 °C) were monitored from 1 to 99 °C. The melts were carried out in a 0.1 cm cuvette on samples with protein concentrations ranging from 10 to 60 μM. Unfolding was completely reversible under all conditions and was not affected by protein concentration. GuHCl denaturation experiments were carried out using a Microlab 500 series automated titrator by stepwise mixing of a protein solution with no GuHCl into a protein solution with high GuHCl (6.0−6.5 M). At each step the sample was stirred in a 1 cm path length cuvette for 45 s, and then a CD measurement averaged for 30 s was recorded. These experiments were performed at a protein concentration of 5 μM and a temperature of 25 °C. GuHCl denaturation was shown to be fully reversible for the wild-type domain.

Thermal and chemical denaturation data were fit to standard equations by nonlinear least-squares regression (using the program Igor Pro) assuming a two-state transition. Linear extrapolation of the folded and unfolded baselines was used to calculate the free energy of unfolding (ΔG_u) at points in the transition regions of unfolding curves. Thermal unfolding data was fit to

$$\Delta G_u(T) = \Delta H_u(T_m)[1 - (T/T_m)] + \Delta C_p[T - T_m - T[\ln(T/T_m)]] \quad (1)$$

where $\Delta G_u(T)$ is the free energy of unfolding at a given temperature in the transition region, T, and $\Delta H_u(T_m)$ is the enthalpy of unfolding at the midpoint temperature of the unfolding transition, T_m. ΔC_p is the heat capacity change upon unfolding, and it was held constant at a value of 680 kcal/mol·deg for all fits. This value was determined in our laboratory for the wild-type protein using the method of Pace (*28*) (data not shown). GuHCl unfolding data was fit to

$$\Delta G_u' = \Delta G_u(H_2O) - m[\mathrm{GuHCl}] \quad (2)$$

where $\Delta G_u'$ is the free energy of unfolding at a given concentration of GuHCl in the transition region, $\Delta G_u(H_2O)$ is the free energy of unfolding at zero denaturant, and m is the dependence of ΔG_u on the concentration of GuHCl. The $\Delta G_u(H_2O)$ values for highly destabilized mutants could be less accurate because their unfolding curves have few points in the folded baseline. However, these curves were also fit using average baseline values from other melts, and similar thermodynamic parameters were obtained.

Peptide Binding. Peptide-binding titrations were carried out by stepwise mixing of the target peptide and SH3 domain using a Microlab 500 series automated titrator. After mixing, the sample was equilibrated for 1 min, and then the degree of binding was assessed by measuring the total tryptophan fluorescence above 320 nm (signal averaged for 30 s) using an excitation wavelength of 295 nm. These measurements were performed using a total fluorescence accessory attached to our CD spectrometer. The target peptide sequence used (VSLARRPLPPLP) was previously isolated by phage display and was shown to bind strongly to the Fyn SH3 domain (*29*). For ease of purification, we expressed this sequence in *E.*

[1] Abbreviations: GuHCl, guanidine hydrochloride; CD, circular dichroism; WT, wild type.

coli with the N-terminal domain of λ repressor (residues 1−99) fused to its N terminus and a hexahistidine sequence at its C terminus. This fusion protein vector was constructed using synthetic oligonucleotides and was derived from the λ repressor expression vector, pN^{102LT} (*30*). Because the region of λ repressor used in this fusion protein contains no tryptophans, it was not necessary to cleave the target to perform binding assays. The binding affinity of the target peptide did not appear to be affected when used in this form. The binding data were fit to

$$I = (([L_T] - ((-([P_T] - [L_T] + K_d) + \mathrm{SQRT}(([P_T] - [L_T] + K_d)^2 + 4K_d[L_T]))/2))/[P_T])(I_\infty - I_0) + I_0 \quad (3)$$

where I is the fluorescence intensity at a given total concentration of target peptide, $[L_T]$, and $[P_T]$ is the total concentration of SH3 domain in the reaction (1 μM in these experiments), K_d is the dissociation constant, I_∞ is the fluorescence intensity at saturation, and I_0 is the intensity in the absence of ligand. The equation used here is different from eq 4, which has been used in other publications (*8*):

$$I = (I_0 + I_\infty([L]/K_d))/(1 + ([L]/K_d) \quad (4)$$

where [L] is the concentration of free ligand. Equation 4 does not give correct K_d values when the K_d is close to $[P_T]$ because [L] cannot be determined directly from the data. When the K_d is higher, [L] can be approximated as being equal to $[L_T]$. Using eq 3 avoided any approximations and allowed us to fit the data with I_0, I_∞, and K_d as free parameters.

RESULTS

Thermodynamic Characterization of the Fyn SH3 Domain. cDNA encoding the SH3 domain of the Fyn tyrosine kinase (residues 85−142) was cloned and expressed in *E. coli*. The sequence of the Fyn SH3 domain is shown in Figure 1B with the residue numbering system used in this work. The protein characterized in this work comprises the SH3 domain plus a short N-terminal and a longer C-terminal tail (see Experimental Procedures). The CD spectrum of this domain is similar to that seen for some other SH3 domains (*8, 9, 31*), and a distinct difference is seen between the spectra of the domain under native and denaturing conditions (Figure 2). The folded CD spectra of Fyn and a number of other SH3 domains are unusual in that they display positive ellipticity with a maximum at ~220 nm rather than negative ellipticity with a minimum at ~217 nm that would be expected for a β-sheet protein. This phenomenon may be a result of the high aromatic residue content of SH3 domains (*32*) or could be due to β-turn conformations, which are known to display positive ellipticity in the 220 nm range (*33, 34*).

In this work, the unfolding of the Fyn SH3 domain was assessed by monitoring the change in CD signal at 220 nm. The chemical and thermal denaturation curves produced in this way were cooperative and were fit by a two-state model (Figure 3). With a T_m of 80.1 °C and a ΔG_u of 4.99 kcal/mol (Table 1), this SH3 domain is similar in stability to most other SH3 domains on which folding studies have been performed (*8, 10, 31*). A kinetic and thermodynamic study by Plaxco et al. (*11*) demonstrated that the Fyn SH3 domain unfolds in a reversible two-state manner with a ΔG_u of 6.0 kcal/mol. The domain used in this study comprised a larger segment of the Fyn protein (67 residues versus 58), and the N and C termini were different from those in our construct, which could account for the higher stability observed.

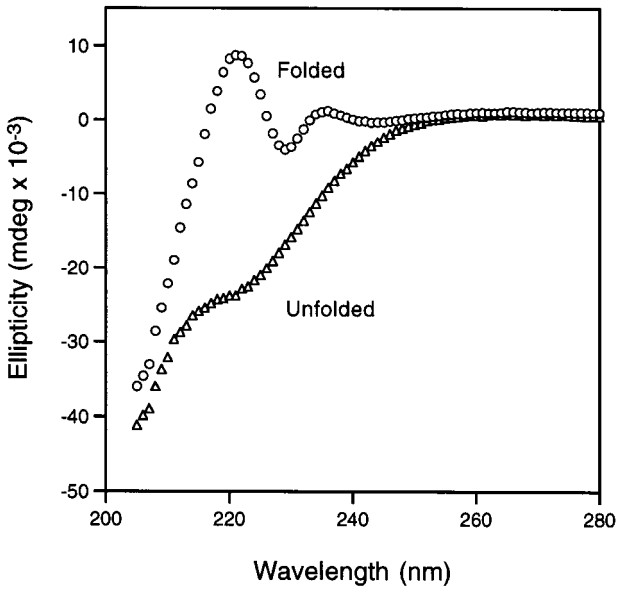

FIGURE 2: CD spectra of the folded and unfolded WT Fyn SH3 domain. These spectra were obtained from a 50 μM sample of protein in a 1 mm cuvette. The folded and unfolded samples were at 25 and 99 °C, respectively.

Mutagenesis of the Glu24 and Ser41 Positions. To investigate the roles of the Glu24 and Ser41 positions in determining the stability of the Fyn SH3 domain, we mutagenized them using randomized oligonucleotides. The objectives of the mutagenesis strategy were to produce substitutions with a variety of residue types at each position and to include all of the residues that are frequently observed at the two positions in the SH3 family. Ultimately, 14 different Glu24 mutants and 13 Ser41 mutants were isolated and sequenced. Typical GuHCl and thermal denaturation curves for WT and some mutants are shown in Figure 4. Table 1 shows the thermodynamic parameters derived for all mutants examined in this study.

Stability of Mutants Substituted at the Glu24 Position. Residue substitutions at the Glu24 position display a wide range of thermodynamic stabilities (Table 1). In general, polar residues appear to be better tolerated than nonpolar residues at this position because substitutions with Asp, Thr, Gln, His, and Lys all have T_m values >70 °C, and the Ser substitution has a T_m just under 70 °C. These residues may be able to maintain one or both of the hydrogen bonds that are made by the WT residue. However, neither of these hydrogen bonds can be crucial for the stability of the domain because the substitution with Ala is still quite stable with a T_m of 72.5 °C and a ΔG_u of 3.72 kcal/mol. The E24D mutant displays a significant increase in stability, especially in its ΔG_u, which is increased by 0.68 kcal/mol. This increase in stability is surprising given the conservative nature of the substitution.

The large degree of destabilization resulting from substitutions with Ile, Leu, or Arg is particularly striking. These mutants display T_m values reduced by >25 °C and $\Delta\Delta G$ values of >2.5 kcal/mol. The effects of the highly desta-

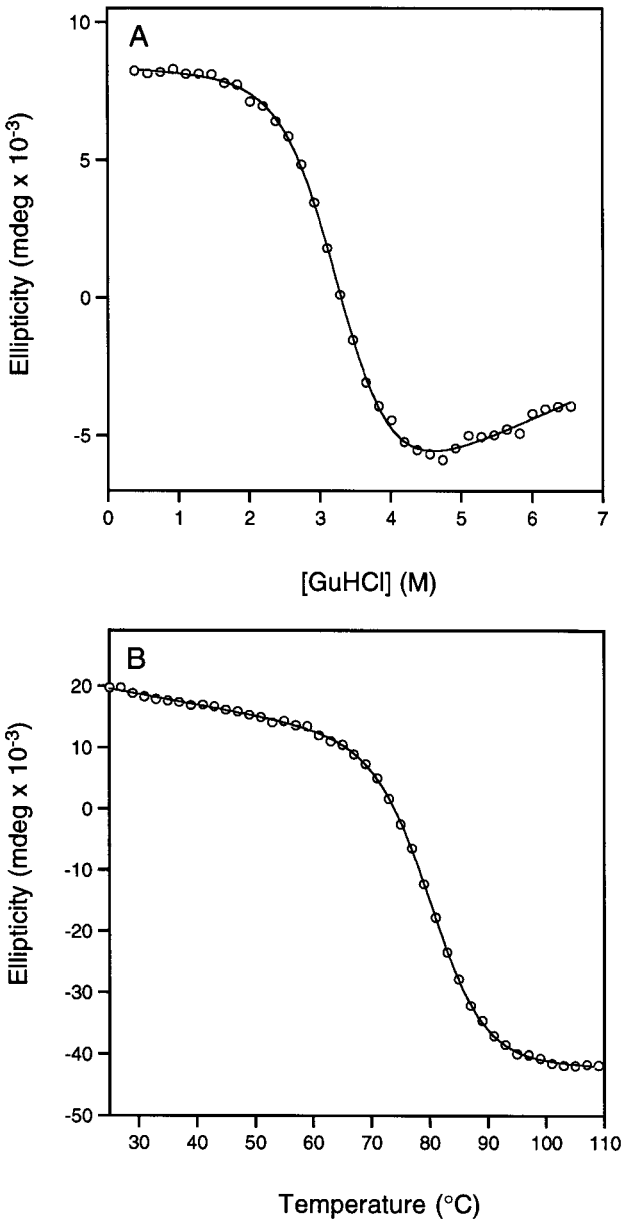

FIGURE 3: Unfolding of the WT Fyn SH3 domain induced by GuHCl (A) and heat (B). The lines joining the points in each graph are the theoretical fits to the data. GuHCl denaturation was performed on a 5 μM protein sample at 25 °C in a 1 cm cuvette. Thermal denaturation was performed on a 100 μM protein sample in a 1 mm cuvette.

bilizing substitutions are difficult to explain because other similar substitutions are not as destabilizing. For example, the substitution of Glu24 with Lys, which is also a large positively charged residue, has a relatively small effect compared to substitution with Arg. Similarly, substitutions with Val and Phe are significantly less destabilizing than substitutions with Ile and Leu.

Stability of Mutants Substituted at the Ser41 Position. The Ser41 position is clearly more tolerant to substitution than the Glu24 position (Table 1). A variety of residue types, including large hydrophobic (Leu, Ile, Val), large charged (Lys, Arg, His), small hydrophobic (Ala), and small polar (Asn), can be substituted at this position without lowering the T_m below 70 °C or reducing the ΔG_u by >1.5 kcal/mol. Furthermore, the lowest T_m observed for a Ser41 mutant is just under 60 °C, whereas Glu24 mutants had T_m values as low as 45 °C. The most destabilized Ser41 mutants can also be rationalized more easily than the most destabilized Glu24 mutants. Gly may be particularly destabilizing at the Ser41 position because it lies within a β-sheet where Gly is known to be energetically unfavorable (*35*, *36*). The instability of the S41D mutant is likely due to repulsion resulting from close positioning of the negatively charged Asp residue to Glu24. An inability to properly pack the bulky aromatic side chains of Tyr and Phe into the hydrophobic core of the domain may explain the low stability of the S41Y and S41F mutants.

The conservative substitution of Ser41 with Thr was surprisingly destabilizing, causing an almost 15 °C drop in T_m and a 1.7 kcal/mol reduction in ΔG_u. Because the replacement of Ser41 with Val, a residue possessing size and shape similar to those of Thr, is not as highly destabilizing, the instability of S41T is probably caused by the inability of its buried hydroxyl group to form a hydrogen bond with Glu24. Another striking result was the relative stability of substitutions of Ser41 with the large hydrophobic residues, Leu, Ile, and Val. These substitutions might have been expected to destabilize the domain substantially because they could force the Glu24 side chain into a completely nonpolar environment. The stability of these mutants, particularly S41L, which has a $\Delta\Delta G_u$ of only 0.4 kcal/mol, suggests that the Glu24 side chain is able to shift position and escape the unfavorable hydrophobic environment created by these substitutions. There may also be adjustments in the hydrophobic core allowing the substituted hydrophobic residues to make favorable packing interactions.

Correlations of the Mutant Thermodynamic Parameters. Because denaturant-induced and thermal unfolding experiments were performed on each mutant, it was important to determine how well the values derived from these experiments were correlated. Figure 5A shows a plot of the ΔG_u of each single-site mutant versus its T_m. It can be seen that this plot is linear with a high correlation coefficient (r value) of 0.95. The only significant outliers are the E24D and S41G mutants. Both of these have higher ΔG_u values than would be expected from their T_m values, behavior that could result from some alteration in the denatured state of these mutants which would lead to a change in their ΔC_p values. The correlation between ΔG_u and T_m for most mutants is consistent with the observation that their m values were close to that of WT (Table 1). The m value of a mutant may change significantly if its denatured state has an altered structure (*37*).

A significant decrease in the ΔH_u of a mutant can indicate that its structure is not as well-packed as WT. Because the ΔH_u values in Table 1 are calculated at the T_m of each mutant, a plot of these values versus T_m should produce a straight line. As can be seen in Figure 5B, this plot is indeed linear with an r value of 0.97. The slope of this plot is 829 kcal/mol·deg, which is, as expected, close to the ΔC_p that we calculated for this domain (see Experimental Procedures).

Functional Effects of Substitutions at the Glu24 and Ser41 Positions. Peptide-binding assays were performed to determine whether mutants still possessed the wild-type native conformation, because substitutions causing a loss of conformational specificity would be expected to display decreased binding affinity. Because Glu24 and Ser41 lie far from the binding surface of the domain (Figure 1A),

Table 1: Effects of Glu24 and Ser41 Substitutions on Stability and Function

protein	T_m (°C)	ΔH_u (T_m) (kcal/mol)	ΔG_u (H_2O) (kcal/mol)	m (kcal/mol·M)	K_d (μM)
WT	80.1[a] ± 0.7	51.4 ± 1.0	4.99[b] ± 0.13	1.50 ± 0.10	0.43[c] ± 0.08
E24D	82.2	53.2	5.67	1.57	0.30
E24T	74.4	46.0	4.39	1.72	0.45
E24Q	74.1	45.6	4.25	1.72	0.58
E24H	73.6	44.9	4.00	1.63	0.34
E24K	73.5	43.7	4.02	1.68	0.54
E24A	72.5	41.7	3.72	1.53	0.83
E24S	67.8	40.0	3.35	1.64	0.65
E24F	66.6	36.6	3.36	1.81	0.56
E24V	65.8	36.0	2.99	1.69	0.37
E24P	64.2	36.6	3.28	1.77	0.74
E24Y	61.2	29.8	2.59	1.98	0.51
E24I	55.2	28.8	2.28	1.92	1.10
E24L	53.0	26.5	1.97	1.83	0.54
E24R	45.6	22.8	1.61	1.88	0.79
S41K	85.5	53.3	5.10	1.78	0.39
S41R	82.6	50.7	4.81	1.73	0.56
S41N	77.5	41.4	4.26	1.55	0.69
S41L	75.6	44.8	4.62	1.70	0.40
S41H	74.8	43.6	3.74	1.66	0.54
S41A	73.9	42.1	4.28	1.60	0.18
S41I	73.3	42.0	3.58	1.47	0.62
S41V	70.3	39.0	3.89	1.60	0.44
S41T	65.6	33.9	3.34	1.72	0.43
S41F	61.4	28.9	3.09	1.79	0.82
S41Y	61.4	30.6	2.72	1.92	0.73
S41D	59.4	30.8	2.97	1.76	0.19
S41G	58.6	27.4	3.31	2.16	0.26
E24A/S41A	68.2	39.7	3.51	1.60	0.76
E24A/S41K	71.4	39.1	3.82	1.61	1.03
E24A/S41L	77.1	45.1	4.56	1.65	1.34
E24L/S41A	51.4	24.3	2.28	1.94	2.03
E24R/S41D	70.8	40.3	3.65	2.02	0.84

[a] Averaged from eight experiments performed at protein concentrations from 10 to 100 μM. [b] Averaged from five experiments. [c] Averaged from three experiments.

substitutions were not expected to affect binding affinity directly.

Peptide binding was measured by monitoring the increase in SH3 domain fluorescence observed as the concentration of target peptide was increased. This increase in fluorescence is caused by the burial of Trp 36, which lies in the peptide-binding groove. We used a target peptide sequence (VSLARRPLPPLP) that had been previously isolated by phage display and was shown to bind strongly to the Fyn SH3 domain (*29*). Although our binding assays were carried out using a fusion between the target peptide and the N-terminal domain of λ repressor, our average measured K_d for WT was 0.43 μM, which is close to the reported K_d of 0.60 μM for the interaction of this peptide with the Fyn SH3 domain (*29*).

The measured K_d values for all of the mutants studied are shown in Table 1. Most of the K_d values show insignificant deviations from the WT value. Even the largest deviation (E24I) amounts to less than a 3-fold increase from WT. This magnitude of increase is small compared to other mutants studied in our laboratory which display >100-fold increases in K_d (data not shown). There is also no correlation between the changes in binding affinity and stability.

Contribution of the E24−S41 Interaction to Thermodynamic Stability. The apparent energy of interaction between two side chains (ΔG_{int}) can be calculated as the difference in the free energy of unfolding upon mutating one of the partners in the presence or absence of the other partner (*38*). To carry out this "thermodynamic cycle" analysis for the Glu24−Ser41 interaction, a double mutant was constructed in which both residues were mutated to Ala (Table 1). Determining the thermodynamic parameters for this mutant allowed us to construct the thermodynamic cycle shown in Figure 6A. It can be seen that substitution of Ser41 with Ala in the absence of its Glu24 hydrogen bonding partner (E24A/S41 to E24A/S41A) has a 0.21 kcal/mol energetic penalty. This indicates that Ser41 contributes 0.21 kcal/mol to the stability of the domain independently of its interaction with Glu24. Mutation of Glu24 causes a 0.77 kcal/mol decrease in stability in the absence of the Ser41 side chain (E24/S41A to E24A/S41A). These data indicate that the effects of the Glu24 and Ser41 mutations are not additive and that the apparent ΔG_{int} of the Glu24 and Ser41 side chains is 0.50 kcal/mol (this value is obtained by finding the difference between opposing $\Delta\Delta G$ values in the thermodynamic cycle).

Formation of a Salt Bridge between the E24 and S41 Positions. The WT level of stability displayed by the S41K and S41R mutants (Table 1) suggested that these positively charged residues might be forming a salt bridge with Glu24 in place of the hydrogen bond formed with the WT Ser residue. To assess the energetic contribution of the potential salt bridge formed between Glu24 and Lys at the Ser41 position, we performed a thermodynamic cycle analysis as described above. One additional double mutant (E24A/

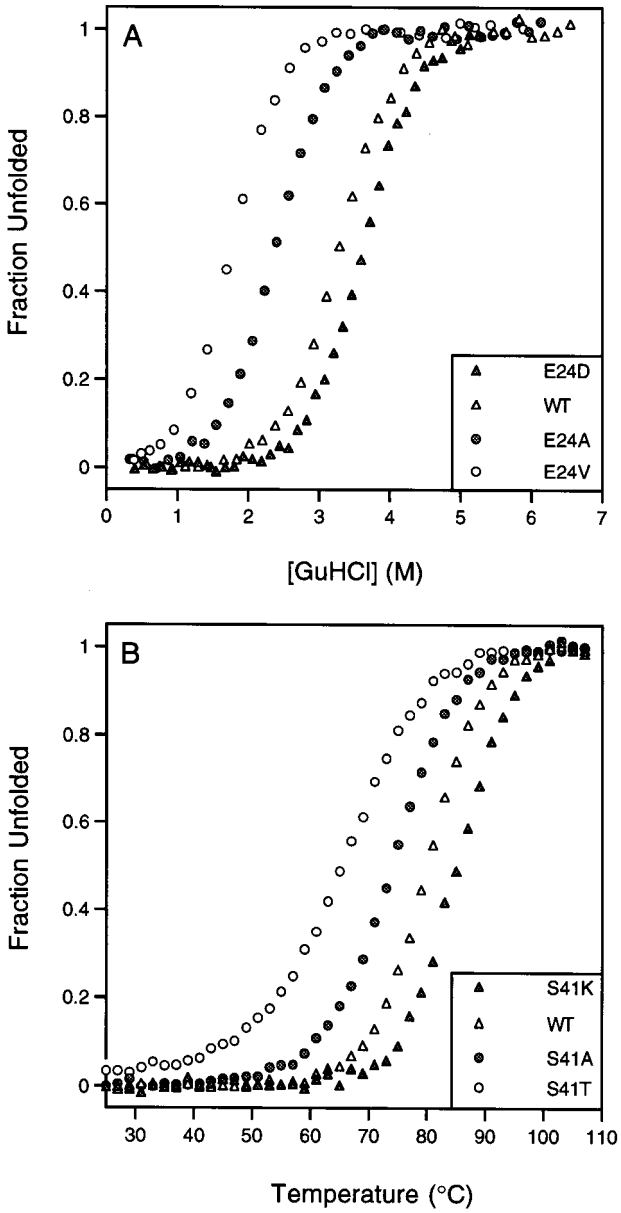

FIGURE 4: Representative GuHCl denaturation (A) and thermal denaturation (B) experiments performed on Glu24 and Ser41 mutants. Fraction unfolded values were calculated from the folded and unfolded baselines used in fitting the curves.

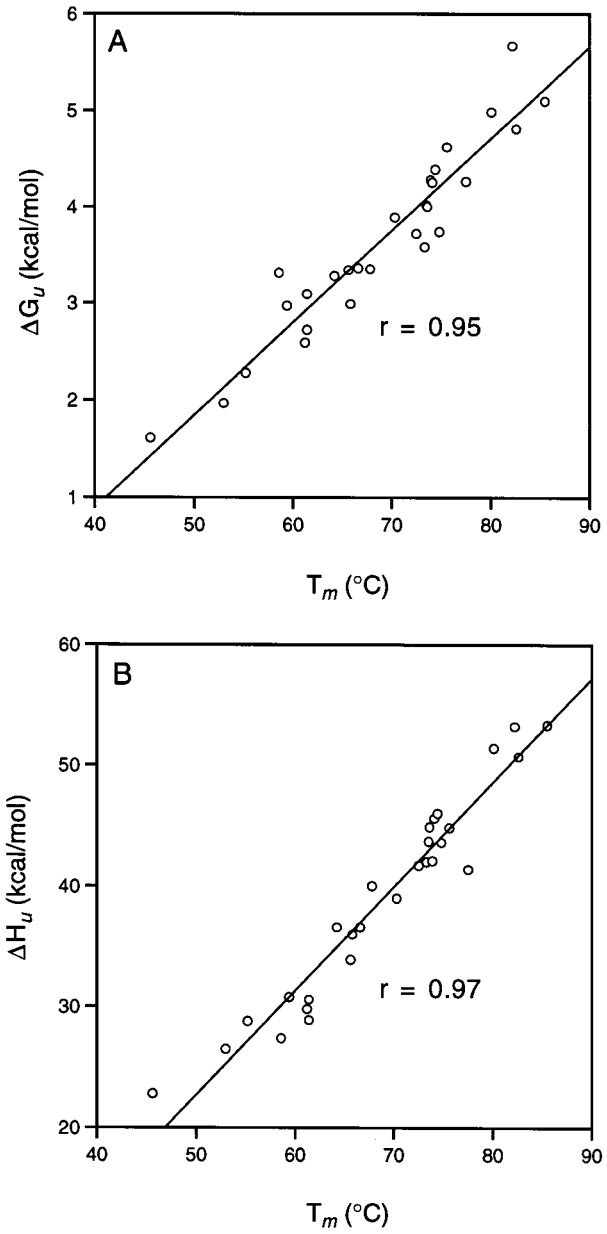

FIGURE 5: Correlations of thermodynamic values of Glu24 and Ser41 mutants. ΔG_u versus T_m (A) and ΔH_u versus T_m (B) are plotted from the values shown in Table 1.

S41K) was constructed for this purpose (Table 1). The thermodynamic cycle in Figure 6B indicates that Glu24 and S41K are indeed interacting and that their interaction contributes 0.51 kcal/mol to the stability of the domain. Strikingly, the energy of this putative salt bridge is almost exactly the same as that of the buried polar interaction between the wild-type residues.

To further investigate the potential for salt bridge formation between the Glu24 and Ser41 positions, the E24R/S41D double mutant was constructed. The E24R and S41D mutants are highly destabilized, with $\Delta\Delta G$ values of 3.5 and 2.0 kcal/mol, respectively. However, when these two substitutions were combined, the double mutant was considerably more stable than either of the single mutants (Table 1). This stabilization can be accounted for if the Arg and Asp at the Glu24 and Ser41 positions are able to form a salt bridge.

Attempted Replacement of the Glu24−Ser41 Interaction with a Hydrophobic Interaction. Because a previous study involving bacteriophage P22 arc repressor showed that replacement of a buried salt bridge with a hydrophobic interaction could dramatically increase protein stability (*20*), we wanted to determine whether the Glu24−Ser41 interaction could be replaced in a similar manner. For this reason, we constructed 21 mutants that each possessed a different combination of hydrophobic residues at the Glu24 and Ser41 positions. Of the 20 mutants constructed, 18 accumulated at such low steady-state levels within *E. coli* that it was impossible to purify them (Table 2). In our characterization of single-site mutants in this study and in studies on many other mutants isolated in our laboratory, we have found that mutants that do not accumulate in *E. coli* at levels high enough for purification invariably possess T_m values below 60 °C (when using our *trc*-promoter-based expression system). To purify mutants with lower stability, their genes

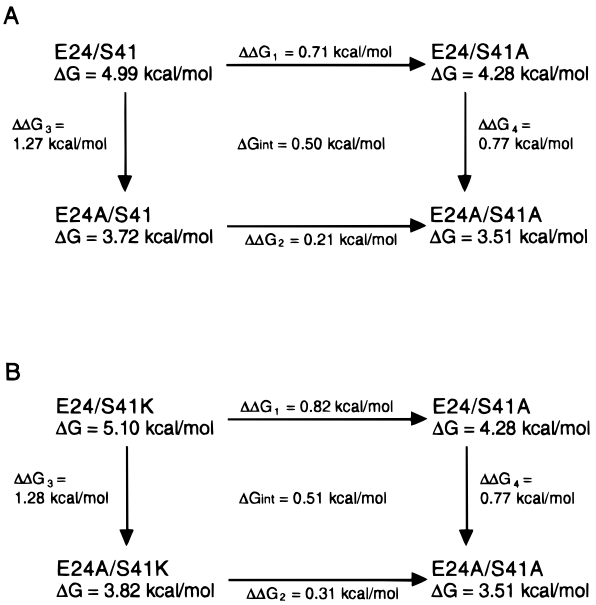

FIGURE 6: Thermodynamic double-mutant cycles for (A) the Glu24−Ser41 interaction and (B) the Glu24−Lys41 interaction.

Table 2: Unstable Hydrophobic Double Mutants

E24L/S41L	E24L/S41V	E24L/S41I	E24L/S41Y	E24L/S41A
E24F/S41L	E24F/S41V	E24F/S41I	E24F/S41Y	E24F/S41A
E24V/S41L	E24V/S41V	E24V/S41I	E24V/S41Y	E24V/S41F
E24I/S41L	E24Y/S41V	E24Y/S41Y		

must be expressed from a T7 promoter. Thus, we can conclude that all of the mutants listed in Table 2 are likely to possess T_m values below 60 °C and are certainly not close to WT in stability. The only hydrophobic combinations that had moderate stability were E24A/S41A and E24A/S41L (Table 1). The increased stability of the E24A/S41L double mutant compared to the E24A mutant (+0.8 kcal/mol) is notable, suggesting that the S41L substitution is able to compensate for the destabilization caused by the E24A substitution. By expression from a T7-promoter-based vector, we were able to characterize the E24L/S41A double mutant. We expected that this mutant might be more stable than the E24L single mutant because the potentially highly buried hydroxyl group of Ser41 would be removed. As is shown in Table 1, however, the E24L/S41A mutant was not significantly more stable than E24L.

DISCUSSION

Importance of the Glu24−Ser41 Interaction. The 0.5 kcal/mol contribution of the Glu24−Ser41 interaction to the stability of the Fyn SH3 domain (Figure 6A) is low compared to values obtained in previous studies of internal hydrogen bond networks. In staphylococcal nuclease and ribonuclease T1, side-chain hydrogen bonds have been estimated to contribute ~1.5 kcal/mol on average (*39*, *40*). Other similar buried interactions in λ repressor and barnase were found to contribute 1.2 and 1.9 kcal/mol, respectively, to the stability of these proteins (*19*, *41*). The low ΔG_{int} value obtained from our thermodynamic cycle may be the result of structural adjustments occurring in the single and double Ala mutants, which result in new favorable interactions being formed. The tolerance of the Glu24 and Ser41 positions to substitutions by a variety of different residue types supports the hypothesis that this region is capable of structural adjustment without loss of stability.

The Glu24−Ser41 interaction does not appear to be important for maintaining the conformational specificity of this domain. Although some of the substitutions at these positions were highly destabilizing, none caused a significant loss of peptide binding activity, as would be expected if mutants were unable to fold into the unique native conformation. Another hallmark of non-native or "molten globule"-like protein structures is lowered enthalpies of unfolding (*42*). Figure 5B shows that none of the Glu24 or Ser41 mutants displayed unexpectedly low ΔH_u values. These mutants also did not display altered tryptophan fluorescence emission spectra (data not shown), which could indicate poorly packed hydrophobic cores. In other studies, buried polar interactions between side chains have been suggested to be important for maintaining conformational specificity (*21*, *23*). The results for this particular buried polar interaction do not support this hypothesis. Because the Glu24−Ser41 interaction is one of only two potential side chain−side chain buried polar interactions in the protein, this class of interactions is unlikely to be crucial in the maintenance of the conformational specificity of the Fyn SH3 domain.

Replacement of the Glu24−Ser41 Interaction. Both theoretical and experimental works have suggested that replacement of a buried polar interaction by a hydrophobic interaction can increase protein stability (*20*, *21*). Although we tested 21 different combinations of hydrophobic residues at the Glu24 and Ser41 positions, none resulted in stability greater than that of WT. As discussed below, the Glu24 position appears to have a low tolerance to hydrophobic substitution that is unrelated to the identity of the residue at the Ser41 position. Our failure to find double hydrophobic mutants with increased stability is likely due to this property of the Glu24 position; thus, this study does not nullify the principle of hydrophobic replacement of buried polar interactions as a method to increase protein stability. It does, however, raise a cautionary note because any position of interest in a given protein could be similar to Glu24 and be intolerant to hydrophobic substitution.

Our results do indicate that the interaction between Glu24 and Ser41 can be replaced by a salt bridge, as in the S41K, S41R, and E24R/S41D mutants. Interestingly, none of these replacements increased the stability of the domain, and the interaction energy between Glu24 and Lys at the Ser41 position was the same as that between Glu24 and Ser41. Any increase in stability that might have been expected from the electrostatic interaction between Glu and Lys may be offset by the increased energetic penalty of desolvating the Lys residue. These results suggest that the replacement of a charged-polar interaction with a salt bridge may not prove to be a generally useful method for increasing the stability of proteins.

Value of the Three-Dimensional Structure for Predicting the Stability Effects of Amino Acid Substitutions. The results of this study underscore the difficulty in predicting the effects of amino acid substitutions on protein stability even when the three-dimensional structure of the protein in question is known. In the X-ray crystal structure of the Fyn SH3 domain, the roles of Glu24 and Ser41 in stabilizing the tertiary structure appear to be similar because the side chains of both residues are highly buried and a hydrogen bond is

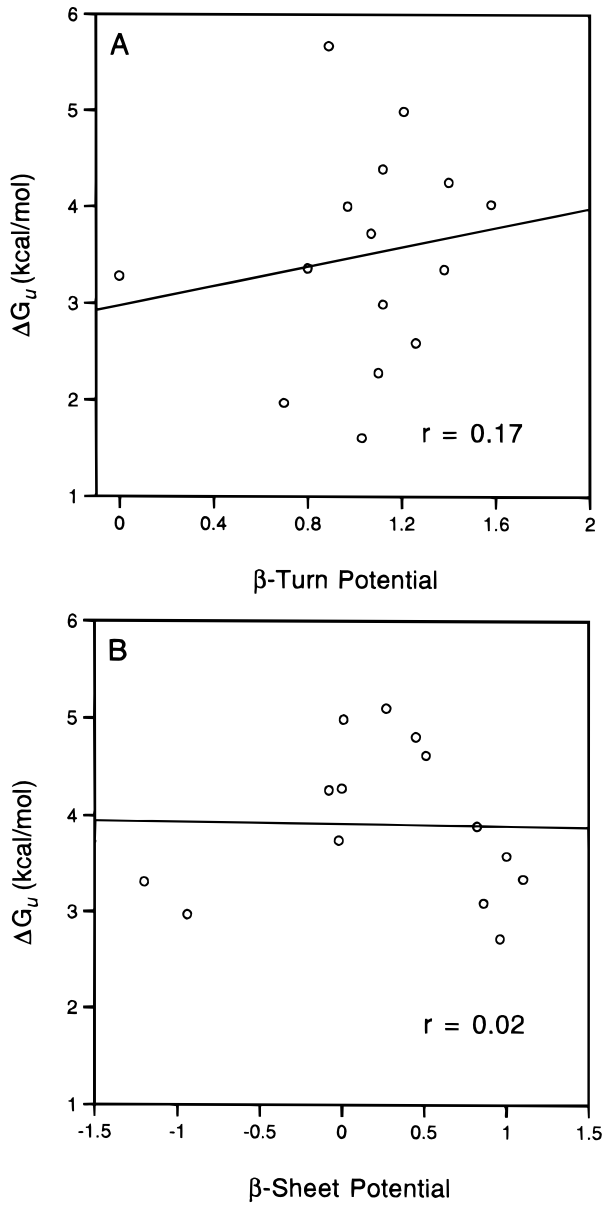

FIGURE 7: (A) ΔG_u values for Glu24 mutants plotted against turn forming potentials for the substituted residues at the $i+3$ position of a type II β-turn (*43*). (B) ΔG_u values for Ser41 mutants plotted against experimentally determined β-sheet propensities for the substituted residues (*35*).

formed between them. The expectation that these positions would, thus, be equally sensitive to mutations is contradicted by the data in Table 1, which shows that Glu24 substitutions were generally more destabilizing than Ser41 substitutions. One could argue that the Glu24 and Ser41 positions display different tolerances for substitutions because they are found in different local conformations: Glu24 lies in a β-turn and Ser41 is part of a β-sheet. If this were true, then some correlation between the effects of mutations at these positions and the propensities for the substituted residue to occupy the local environment would be expected. Figure 7A shows that there is no correlation between the stabilities of substitutions at the Glu24 position and the β-turn-forming potentials of the substituted residues derived from a large structural database study (*43*). Similarly, experimentally determined β-sheet-forming propensities of residues substituted at the Ser41 position (*35*) show no correlation with the stabilities of the Ser41 mutants (Figure 7B). There is also no correlation between mutant stability and either hydrophobicity or side-chain volume of substituted residues.

The divergent effects of Leu and Ile substitutions at the Glu24 and Ser41 positions exemplify the problems inherent in predicting the stabilities of mutants. The substitution of Ser41 with Leu and Ile results in only a small decrease in stability, implying that this region of the protein is able to adjust, thereby allowing the introduction of larger side chains. The buried O_δ atom of Glu24 that interacts with Ser41 in the WT structure must also be able to change position to avoid being buried. In contrast, substituting Glu24 with Ile or Leu destabilizes the domain drastically. This effect is unlikely to be due to burial of the hydroxyl group of Ser41 by the large hydrophobic groups introduced at the Glu24 position because the E24L/S41A double mutant is no more stable than E24L (Table 1). When these substitutions were examined using a computer modeling program (INSIGHT II), it appeared that a number of Leu or Ile rotamers could be placed at the Glu24 position without steric clashes, but there were no possible ways to fit Leu or Ile at the Ser41 position. Thus, a definitive prediction or explanation of the effects of these substitutions from examination of the wild-type structure is not possible. A general conclusion at this point is that the region around the Ser41 position is able to adjust more easily when various residues are substituted. Residues substituted here with significant aliphatic regions (e.g., Leu, Ile, Arg, Lys) may be able to make compensatory favorable packing interactions with nearby hydrophobic core residues (e.g., Phe26).

Value of Sequence Alignment for Predicting the Stability Effects of Amino Acid Substitutions. Because examination of the three-dimensional structure of this domain did not provide accurate predictions of the effects of mutations at the Glu24 and Ser41 positions, we wondered if the analysis of sequence alignment data might be more informative. We have constructed an alignment of 370 different SH3 domain sequences collected by exhaustive BLAST (*44*) searches (manuscript in preparation). The sequences within this alignment were weighted using an established procedure (*45*) to eliminate bias that would occur from over-representation of some SH3 domain subfamilies. The "weighted" frequency of occurrence of residues (percent conservation) at the Glu24 and Ser41 positions along with the stability parameters of our mutants are shown in Table 3. The residues substituted at these positions can be divided into two groups in terms of conservation: "conservative", where the conservation is $\geq$4%, and "non-conservative", where the conservation is $\leq$1%. Table 3 shows that the 13 conservative substitutions have an average T_m of 73.5 °C and ΔG_u of 4.05 kcal/mol, whereas the 14 nonconservative substitutions have an average T_m of 63.9 °C and ΔG_u of 3.18 kcal/mol. In 10 of 13 cases, substitutions with conservative substitutions had T_m values $>$70 °C. The large difference between the stabilities of the conservative and nonconservative substitutions indicates that there may be a significant correlation between conservation and stability. Figure 8 shows a plot of ΔG_u versus percent conservation for all substitutions with residues that occurred at the Glu24 and Ser41 positions with a frequency above zero (the wild-type residues were also included). Due to the asymptotic nature of the data (i.e., the stability cannot greatly exceed that of wild type and the conservation cannot

Table 3: Conservation Observed in the SH3 Domain Family Compared to the Stability of Mutants

protein	conservation[a] %	T_m (°C)	ΔG_u (H_2O) (kcal/mol)
E24D	59	82.2	5.67
E24 (WT)	26	80.1	4.99
E24A	5.8	72.5	3.72
E24Q	4.7	74.1	4.25
E24V	1.1	65.8	2.99
E24T	0.57	74.4	4.39
E24K	0.48	73.5	4.02
E24S	0.47	67.8	3.35
E24H	0	73.6	4.00
E24F	0	66.6	3.36
E24P	0	64.2	3.28
E24Y	0	61.2	2.59
E24I	0	55.2	2.28
E24L	0	53.0	1.97
E24R	0	45.6	1.61
S41L	19	75.6	4.62
S41 (WT)	10	80.1	4.99
S41N	11	77.5	4.26
S41V	9.4	70.3	3.89
S41R	7.9	82.6	4.81
S41D	6.4	59.4	2.97
S41K	6.1	85.5	5.10
S41I	4.5	73.3	3.58
S41H	4.4	74.8	3.74
S41T	4.4	65.6	3.34
S41Y	3.8	61.4	2.72
S41F	1.4	61.4	3.09
S41G	0.60	58.6	3.31
S41A	0.47	73.9	4.28
averages	≥4%	73.5	4.05
	≤1%	63.9	3.18

[a] Refers to the weighted frequency with which each residue occurs at the given position in our sequence alignment of SH3 domains.

be lower than zero), we found that a logarithmic plot could best describe the relationship between these two parameters. As shown in Figure 8, all of the data can be fit with a correlation coefficient of 0.52. This fit allowed the identification of five outliers (points that were off the line by >20%). Fitting the data without these outliers results in a much higher correlation coefficient ($r = 0.79$).

Some of the unexpected results in this study could be predicted from the conservation data. The Glu24 position was found to be generally more sensitive to substitutions than the Ser41 position in that the average T_m of mutants at this position was lower and a more limited selection of residue types was tolerated. This result is consistent with the alignment data which showed that the Glu24 position is the seventh most conserved in the SH3 domain family (as judged by positional entropy, which is a statistical measure of variability), whereas the Ser41 position is one of the least conserved. The observation that the Glu24−Ser41 interaction does not make a large contribution to the thermodynamic stability of the domain is also consistent with the lack of conservation of this interaction in the SH3 domain family. The conservation data also predict the surprisingly high stability of the S41L mutant (Leu is the most common residue observed at this position) and the contrastingly low stability of Leu and Ile substitutions at the Glu24 position.

There are exceptions to the congruence of conservation and stability effects, some of which are seen as outliers in Figure 8. For example, the S41D and S41Y mutants are less stable than would be expected from their degree of

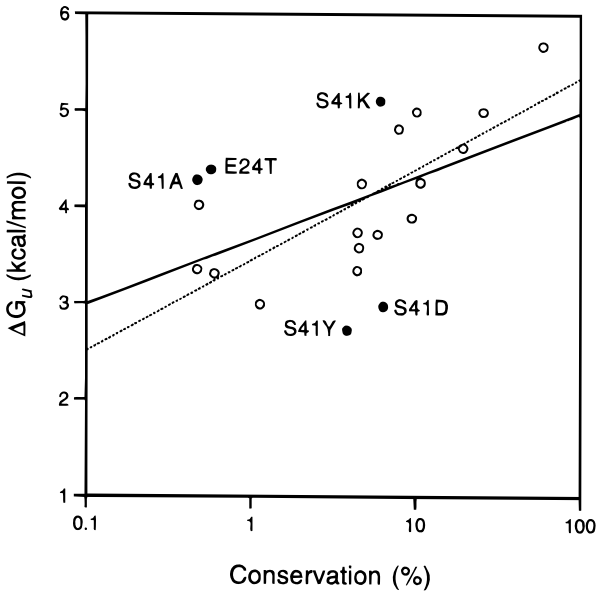

FIGURE 8: Stability versus sequence conservation at the Glu24 and Ser41 positions. The ΔG_u values for each mutant and the wild-type domain were plotted against the conservation values shown in Table 3. A logarithmic curve fit to all data points is shown by the solid line ($r = 0.52$), and a fit with all points, excluding the indicated outliers ($r = 0.79$), is shown by the dotted line. Mutants with conservation values of zero were not included in the plot because they could not be fit by the logarithmic equation. A log scale is used on the *x*-axis to allow all points to be clearly seen.

conservation. The stability of these mutants may be explained by the existence of covariation. When Asp or Tyr appears at the Ser41 position, some other position(s) may generally be a different residue from that found in the Fyn SH3 domain. Constructing an appropriate double mutant, as predicted by covariation analysis, could restore the stability of these mutants. The other class of exceptions to the conservation rule occur when substitution with residues occurring at low frequency in the alignment results in stable functional proteins (e.g., E24K, E24T, E24H, and S41A). The possibility exists that evolution has simply not yet sampled these residues at these positions, but this seems unlikely because single base changes from the most commonly observed residues could produce most of these substitutions. Another possibility is that these mutants, although they appear normal in our in vitro assays, may not function properly in their natural in vivo environment. Tests of these mutants in the context of the complete Fyn tyrosine kinase inside cells would need to be undertaken to verify this hypothesis.

Important Implications of the Relationship between Conservation and Stability. To our knowledge, no previous studies have examined the relationship between conservation in a protein family and the stability effects of amino acid substitutions in a single member of that family. Here we have shown that a significant correlation between conservation and stability is observed at two positions in an SH3 domain. Because evolution selects for function, not high thermodynamic stability, this result would not necessarily be predicted. If SH3 domains with a broad range of stabilities could still carry out their in vivo roles, then conservation and stability might be less correlated. Thus, we suggest that most SH3 domains require a minimum level of stability to function in the cell. This idea is supported by

in vitro folding studies showing that most SH3 domains are similar in stability, with an average ΔG_u of 3.9 ± 0.7 kcal/mol ($n = 8$) (*8*−*10*, *46*, *47*), with the exception of the Drk SH3 domain, which has a ΔG_u of approximately zero (*48*). Another surprising aspect of these results is that the residue frequencies calculated from the whole SH3 domain family, which comprises a widely divergent set of sequences, have predictive value for mutations in one particular domain with a single unique sequence. This suggests that similar interactions stabilize most SH3 domains.

Much research today is focused on the engineering of proteins by site-directed mutagenesis to enhance stability, solubility, or catalytic properties. The success of these projects is often hampered by the difficulty of predicting the effects of amino acid substitutions even when the three-dimensional structure of a protein is known. This study has demonstrated that analysis of sequence alignments can provide a valuable source of information for the prediction of the stability effects of amino acid substitution. At the Glu24 and Ser41 positions, only 46% of substitutions resulted in proteins with T_m values >70 °C. However, 90% of substitutions with residues occurring at a frequency of >5% had T_m values this high. Clearly, the chances of choosing the appropriate residues to substitute in a protein engineering study can be greatly increased by using information from sequence alignments. Furthermore, if one's aim were to produce a protein more stable than wild type, searching for residues in the wild-type sequence that are rarely observed in the family and replacing them with residues that are commonly seen might be a fruitful approach.

ACKNOWLEDGMENT

We thank Julie Forman-Kay and Kevin Plaxco for comments on the manuscript. We also thank Robert Sauer, in whose laboratory these studies were initiated, and Stefan Larson for providing sequence alignment data and comments on the manuscript.

REFERENCES

1. Pawson, T. (1995) *Nature 373*, 573−80.
2. Pawson, T., and Scott, J. D. (1997) *Science 278*, 2075−80.
3. Booker, G. W., Breeze, A. L., Downing, A. K., Panayotou, G., Gout, I., Waterfield, M. D., and Campbell, I. D. (1992) *Nature 358*, 684−7.
4. Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., and Saraste, M. (1992) *Nature 359*, 851−5.
5. Noble, M. E., Musacchio, A., Saraste, M., Courtneidge, S. A., and Wierenga, R. K. (1993) *EMBO J. 12*, 2617−24.
6. Kohda, D., Hatanaka, H., Odaka, M., Mandiyan, V., Ullrich, A., Schlessinger, J., and Inagaki, F. (1993) *Cell 72*, 953−60.
7. Borchert, T. V., Mathieu, M., Zeelen, J. P., Courtneidge, S. A., and Wierenga, R. K. (1994) *FEBS Lett. 341*, 79−85.
8. Lim, W. A., Fox, R. O., and Richards, F. M. (1994) *Protein Sci. 3*, 1261−6.
9. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L. (1994) *Biochemistry 33*, 2142−50.
10. Grantcharova, V. P., and Baker, D. (1997) *Biochemistry 36*, 15685−92.
11. Plaxco, K. W., Guijarro, J. I., Morton, C. J., Pitkeathly, M., Campbell, I. D., and Dobson, C. M. (1998) *Biochemistry 37*, 2529−37.
12. Mayer, B. J., and Eck, M. J. (1995) *Curr. Biol. 5*, 364−7.
13. Stickle, D. F., Presta, L. G., Dill, K. A., and Rose, G. D. (1992) *J. Mol. Biol. 226*, 1143−59.
14. Bowie, J. U., Luthy, R., and Eisenberg, D. (1991) *Science 253*, 164−70.
15. Rashin, A. A., and Honig, B. (1984) *J. Mol. Biol. 173*, 515−21.
16. Barlow, D. J., and Thornton, J. M. (1983) *J. Mol. Biol. 168*, 867−85.
17. Anderson, D. E., Becktel, W. J., and Dahlquist, F. W. (1990) *Biochemistry 29*, 2403−8.
18. Tissot, A. C., Vuilleumier, S., and Fersht, A. R. (1996) *Biochemistry 35*, 6786−94.
19. Marqusee, S., and Sauer, R. T. (1994) *Protein Sci. 3*, 2217−25.
20. Waldburger, C. D., Schildbach, J. F., and Sauer, R. T. (1995) *Nat. Struct. Biol. 2*, 122−8.
21. Hendsch, Z. S., and Tidor, B. (1994) *Protein Sci. 3*, 211−26.
22. Betz, S. F., Raleigh, D. P., and DeGrado, W. F. (1993) *Curr. Opin. Struct. Biol. 3*, 601−10.
23. Lumb, K. J., and Kim, P. S. (1995) *Biochemistry 34*, 8642−8.
24. Davidson, A. R., and Sauer, R. T. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2146−50.
25. Landt, O., Grunert, H. P., and Hahn, U. (1990) *Gene 96*, 125−8.
26. Studier, F. W., and Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113−30.
27. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−23.
28. Pace, C. N., and Laurents, D. V. (1989) *Biochemistry 28*, 2520−5.
29. Rickles, R. J., Botfield, M. C., Zhou, X. M., Henry, P. A., Brugge, J. S., and Zoller, M. J. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 10909−13.
30. Parsell, D. A., and Sauer, R. T. (1989) *Genes Dev. 3*, 1226−32.
31. Chen, Y. J., Lin, S. C., Tzeng, S. R., Patel, H. V., Lyu, P. C., and Cheng, J. W. (1996) *Proteins: Struct., Funct., Genet. 26*, 465−71.
32. Chakrabartty, A., Kortemme, T., Padmanabhan, S., and Baldwin, R. L. (1993) *Biochemistry 32*, 5560−5.
33. Chang, C. T., Wu, C. S., and Yang, J. T. (1978) *Anal. Biochem. 91*, 13−31.
34. Shin, H. C., Merutka, G., Waltho, J. P., Wright, P. E., and Dyson, H. J. (1993) *Biochemistry 32*, 6348−55.
35. Minor, D., Jr., and Kim, P. S. (1994) *Nature 367*, 660−3.
36. Smith, C. K., Withka, J. M., and Regan, L. (1994) *Biochemistry 33*, 5510−7.
37. Shortle, D., and Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81−9.
38. Horovitz, A., and Fersht, A. R. (1990) *J. Mol. Biol. 214*, 613−7.
39. Shirley, B. A., Stanssens, P., Hahn, U., and Pace, C. N. (1992) *Biochemistry 31*, 725−32.
40. Byrne, M. P., Manuel, R. L., Lowe, L. G., and Stites, W. E. (1995) *Biochemistry 34*, 13949−60.
41. Chen, Y. W., Fersht, A. R., and Henrick, K. (1993) *J. Mol. Biol. 234*, 1158−70.
42. Ptitsyn, O. B. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 243−300, W. H. Freeman and Co., New York.
43. Hutchinson, E. G., and Thornton, J. M. (1994) *Protein Sci. 3*, 2207−16.
44. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−402.
45. Henikoff, S., and Henikoff, J. G. (1994) *J. Mol. Biol. 243*, 574−8.
46. Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D., and Dobson, C. M. (1998) *J. Mol. Biol. 276*, 657−67.
47. Knapp, S., Mattson, P. T., Christova, P., Berndt, K. D., Karshikoff, A., Vihinen, M., Smith, C. I., and Ladenstein, R. (1998) *Proteins: Struct., Funct., Genet. 31*, 309−19.
48. Zhang, O., and Forman-Kay, J. D. (1995) *Biochemistry 34*, 6784−94.
49. Evans, S. V. (1993) *J. Mol. Graph. 11*, 134−8, 127−8.

BI981788P