Hydrophobic Core Manipulations in Ribonuclease T1[†]

Stefan De Vos,*[,‡] Jan Backmann,[§,‡] Martine Prévost,[‖] Jan Steyaert,[‡] and Remy Loris[‡]

Laboratorium voor Ultrastructuur, Vlaams Interuniversitair Instituut voor Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius-Rode, Belgium, and Ingénierie Biomoléculaire CP 165/64, Université Libre de Bruxelles, Av. Franklin Roosevelt 50, B-1050 Brussels, Belgium

Received March 20, 2001; Revised Manuscript Received June 26, 2001

ABSTRACT: Differential scanning calorimetry, urea denaturation, and X-ray crystallography were combined to study the structural and energetic consequences of refilling an engineered cavity in the hydrophobic core of RNase T1 with CH_3, SH, and OH groups. Three valines that cluster together in the major hydrophobic core of T1 were each replaced with Ala, Ser, Thr, and Cys. Compared to the wild-type protein, all these mutants reduce the thermodynamic stability of the enzyme considerably. The relative order of stability at all three positions is as follows: Val > Ala ≈ Thr > Ser. The effect of introducing a sulfhydryl group is more variable. Surprisingly, a Val → Cys mutation in a hydrophobic environment can be as or even more destabilizing than a Val → Ser mutation. Furthermore, our results reveal that the penalty for introducing an OH group into a hydrophobic cavity is roughly the same as the gain obtained from filling the cavity with a CH_3 group. The inverse equivalence of the behavior of hydroxyl and methyl groups seems to be crucial for the unique three-dimensional structure of the proteins. The importance of negative design elements in this context is highlighted.

Understanding how the amino acid sequence of a protein encodes its unique three-dimensional structure remains one of the major challenges of modern structural biology (*1*, *2*). A full understanding of the basic principles that relate folding, stability, and structure cannot be achieved by merely knowing the three-dimensional structures of a representative subset of all proteins encoded within a genome. To achieve this goal, knowledge and understanding of the thermodynamics of protein structure and of folding kinetics are as important.

Site-directed mutagenesis is a powerful tool for probing the basics of the relationship between sequence, structure, and function (*3*). Using various model proteins, the effects of substitutions, insertions, and deletions have been investigated. In particular, cavity-creating mutations have been studied extensively by various research teams (*4*, *5*). Isosteric substitutions whereby CH_3, OH, and SH groups are substituted have also been given attention (*6*−*12*). This is not surprising given the information that can be obtained concerning the hydrophobic effect, the presumed driving force of protein folding (*13*, *14*). However, only few of these studies included crystal structure determinations of the mutants that are involved.

Matthews and co-workers studied a number of Ala → Ser and Val → Thr mutations in T4 lysozyme and concluded that depending on the solvent accessibility of the site, the costs of the substitution varied between 0.2 and 2.8 kcal/mol (*15*). The most destabilizing mutant (V149T) incorporates a water molecule, which provides a hydrogen-bonding partner for the buried OH group. Analysis of the crystal structures of these mutants revealed, however, that in each case the buried hydroxyl group is capable of making at least a weak hydrogen bond with a nearby carbonyl group, making the environment not truly hydrophobic. According to Serrano and co-workers (*11*), replacement of a buried Val with Thr also costs ∼2.5 kcal/mol, which was explained as the cost of losing two hydrogen bonds (*11*). Byrne and co-workers on the other hand studied the effect of a large number of small polar and/or nonpolar substitutions on the stability of staphylococcal nuclease (*8*). Strong conclusions in their study were however difficult to make since the crystal structures of these mutants were not determined. In the study presented here, we systematically probed the effects on structure and stability of the introduction of OH and SH groups into a completely hydrophobic cavity of a globular protein.

Ribonuclease T1 (EC 3.1.27.3) is a small (104 amino acids, ∼11 kDa) monomeric enzyme that has been used as a model protein to study enzymatic catalysis (*16*) and protein stability (*17*, *18*). Ribonuclease T1 consists of a 4.5-turn α-helix that is packed upon a major 5-stranded β-sheet (Figure 1a). A second small 2-stranded β-sheet precedes the helix and connects the N- and C-termini. Two hydrophobic cores can be distinguished, located on opposite sides of the major sheet. The smaller of the two cores is located on the concave side of the major sheet harboring the catalytic site. The larger core has a more extended shape and docks the α-helix and the small β-sheet upon the major β-sheet. The core is

[†] This work was supported by the Vlaams Interuniversitair Instituut voor Biotechnologie and the Fonds voor Wetenschappelijk Onderzoek Vlaanderen. R.L. is a research associate of the Fonds voor Wetenschappelijk Onderzoek Vlaanderen. S.D.V. received a grant from the Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek in de Industrie (IWT).

* To whom correspondence should be addressed: Laboratorium voor Ultrastructuur, Vlaams Interuniversitair Instituut voor Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius-Rode, Belgium. Telephone: 32-2-3590268. Fax: 32-2-3590289. E-mail: stedevos@vub.ac.be.

[‡] Vrije Universiteit Brussel.

[§] Current address: AISE, Square Marie-Louise 49, B-1000 Brussels, Belgium.

[‖] Université Libre de Bruxelles.

10.1021/bi010565n CCC: $20.00
Published on Web 08/02/2001

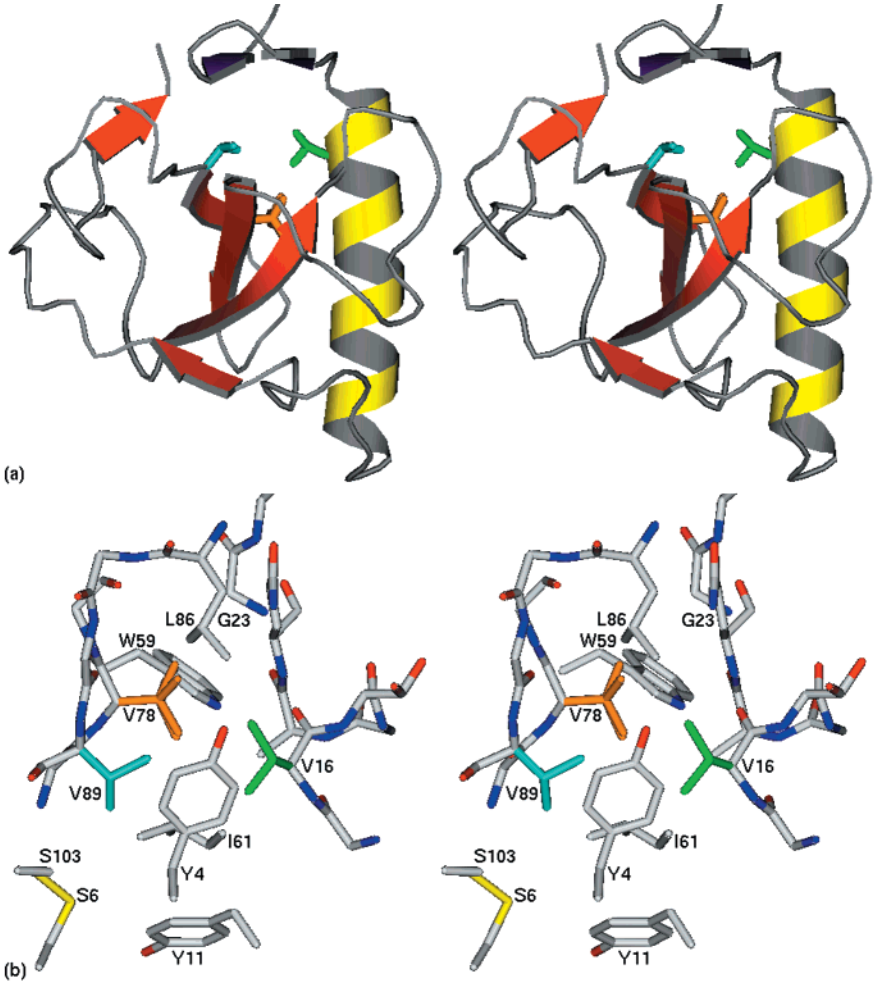

FIGURE 1: Structural context of Val16, Val78, and Val89 in wild-type RNase T1. (a) Stereoview of a ribbon representation of RNase T1. The α-helix is shown in yellow, the major N-terminal β-sheet in red, and the minor β-sheet in purple. The three residues mutated in this study are shown in a ball-and-stick representation: Val16 in green, Val78 in orange, and Val89 in cyan. (b) Ball-and-stick representation of the structural neighborhood of Val16, Val78, and Val89. These three residues are color-coded as in panel a. Other neighboring residues in this hydrophobic core are also shown and labeled.

centered on Cys6, Val16, Gly23, Trp59, Ile61, Val78, and Val89 (Figure 1b). The side chains of these residues are completely or almost completely buried and are surrounded by a layer of partially buried and predominantly hydrophobic residues.

EXPERIMENTAL PROCEDURES

Oligonucleotide-Directed Mutagenesis and Enzyme Purification. The overproduction of recombinant wild-type RNase T1 and the mutants as secretory proteins in *Escherichia coli* has been described previously (*19*). The mutants were constructed with a polymerase chain reaction (PCR[1])-based site-directed mutagenesis technique (*20*). Mutations were identified by DNA sequence determination using an ALF-Express II system (Amersham Pharmacia Biotech). The correct sequence of the entire gene was verified. Enzymes were purified to homogeneity as described previously (*21*), and their molecular weights were confirmed by electrospray mass spectrometry.

Urea Unfolding. Samples were prepared by diluting appropriate volumes of a concentrated protein stock solution and a concentrated urea solution (pH-adjusted, concentration of urea determined by means of a refractometer) into protein-free measurement buffer [30 mM MOPS (pH 7.0)] to have constant protein concentrations and different denaturant concentrations (*22*). The degree of urea-induced denaturation of wild-type RNase T1 and its mutants was observed by fluorescence with an AMINCO-Bowman Series2 luminescence spectrometer (Spectronic Instruments, Rochester, NY), where $\lambda_{exc} = 278$ nm and $\lambda_{em} = 320$ nm. Measurements were taken at 5, 10, 15, 20, and 25 °C to study the temperature dependency of the conformational stability.

The obtained unfolding isotherms were processed applying a linear extrapolation method (*22*) based on the fitting function described by Santoro and Bolen (*23*). It was assumed that the *m* value does not depend on temperature. That is why the *m* value obtained after a first round of fitting was averaged over all measurements. In a second round of fitting, the *m* value was fixed (nonfloatable parameter) to obtain a more reliable value for $\Delta G(H_2O)$. Similar values for $\Delta G(H_2O)$ can be obtained using the mean *m* value, the

[1] Abbreviations: DSC, differential scanning calorimetry; CD, circular dichroism; $\Delta G(H_2O)$, free energy in the absence of denaturant; ΔH_{trs}, calorimetric enthalpy of unfolding at the transition temperature; T_{trs}, temperature at which the fraction of unfolded protein equals $^1/_2$; ΔC_p, change in heat capacity upon unfolding; PCR, polymerase chain reaction; 2′GMP, 2′-guanylic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid.

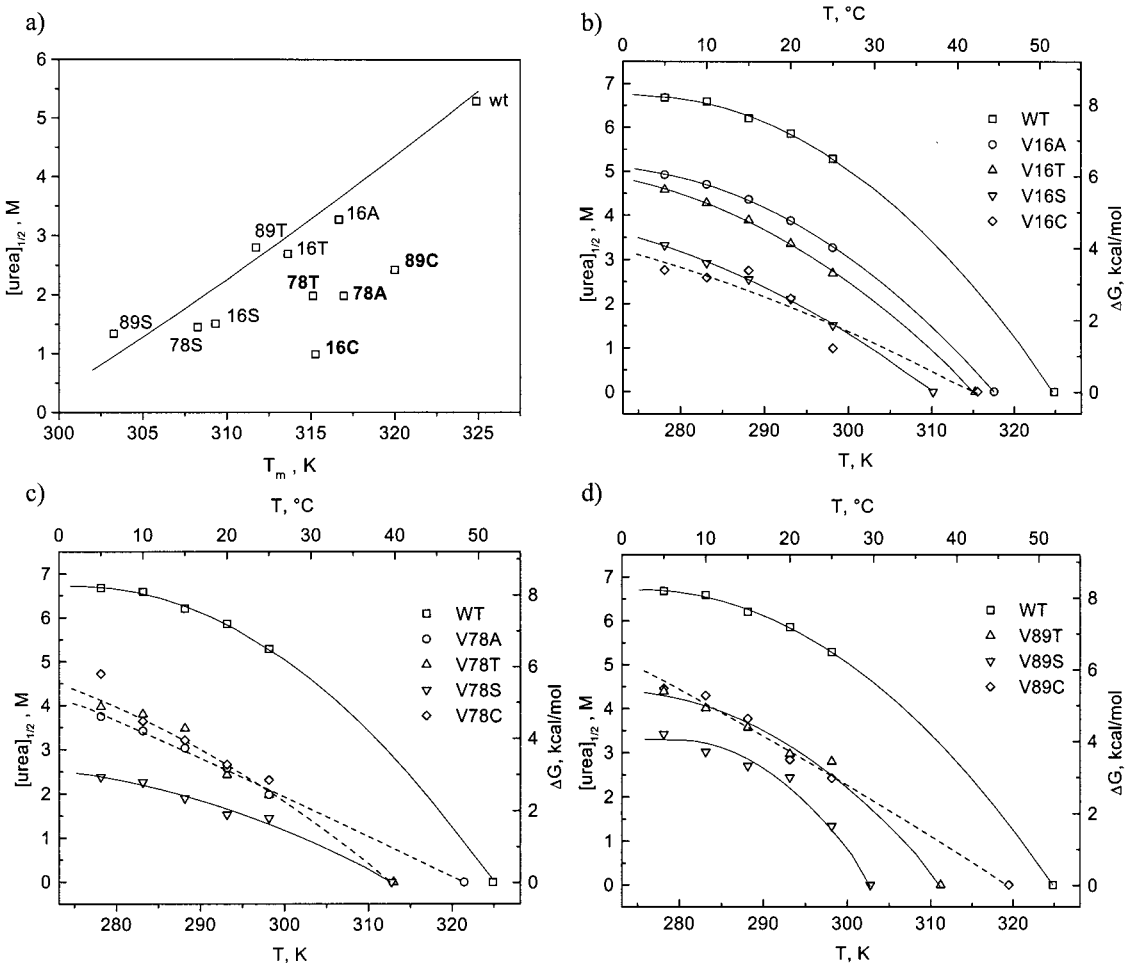

FIGURE 2: (a) Comparison between thermal and urea-induced unfolding. A plot is shown of $[\text{urea}]_{1/2}$ versus T_{trs}, according to eq 3. The black line corresponds to a theoretical curve that predicts this relationship assuming a constant ΔC_p. Data points are labeled; bold labels correspond to experimental measurements that deviate significantly from the theoretical curve, suggesting changes in ΔC_p. (b−d) Temperature dependence of $\Delta G_{unfolding}$ at the three target locations. The plotted points are the values for $\Delta G_{unfolding}$ determined at 5, 10, 15, 20, and 25 °C using urea denaturation as well as the melting temperature (at which $\Delta G_{unfolding} = 0$ for a monomeric protein) derived from CD measurements. In each case, except V78C (where no melting temperature was determined), a fit to the data using eq 2 is shown superimposed.

individual urea half-points ($[\text{urea}]_{1/2}$), and eq 1 (*24*):

$$\Delta G(H_2O) = -m[\text{urea}]_{1/2} \quad (1)$$

Differential Scanning Calorimetry. DSC measurements were performed on an MC-2 scanning calorimeter (MicroCal, Northampton, MA). Thermograms were obtained at a scan rate of 60°/h. The protein solutions for the calorimetric measurements were prepared by dissolving the lyophilized protein in 30 mM MOPS (pH 7.0) and equilibrated by passing them over a gravity-driven disposable gel filtration column. The concentration was determined using UV absorbance and a molar extinction coefficient ϵ_{278} of 19 330 M^{-1} cm^{-1} (*25*). The protein solutions were degassed by stirring them for at least 5 min under vacuum. Two consecutive scans were carried out to observe the extent of reversibility of protein unfolding. For all mutants but the V16C mutant, the experiments revealed a high level of reversibility.

Circular Dichroism. Thermal denaturation experiments using far-UV circular dichroism were carried out with a Jasco J-715 spectropolarimeter using quartz cells with an optical path length of 0.1 cm. Protein concentrations of 0.2 mg/mL in 30 mM MOPS (pH 7.0) were used. The temperature of the samples was controlled during the measurements using a sensor built into the cuvette holder and connected to a Haake N3 circulating bath with an accuracy of 0.1 °C.

Thermodynamic Data Analysis. To correlate the urea and the temperature-induced unfolding experiments, a modified Gibbs−Helmholtz equation (eq 2) has been used. A constant ΔC_p value of 1290 cal mol^{-1} K^{-1} (taken from ref *26*) was assumed throughout the fitting process.

$$\Delta G(T) = \Delta H_{trs}\left(\frac{T_{trs} - T}{T}\right) - \Delta C_p(T_{trs} - T) + T\Delta C_p \ln\left(\frac{T_{trs}}{T}\right) \quad (2)$$

In unfolding analysis, it is often assumed that ΔC_p is temperature-independent (*27*). To validate this assumption, the urea half-points of the wild type and the different mutant proteins are plotted against their melting temperatures (Figure 2a). Equation 3 (*28*), which resulted from a combination of eqs 1 and 2, renders the ideal situation where the assumption of a constant ΔC_p is valid. If ΔC_p does not change as a function of T, the $[\text{urea}]_{1/2}$ is related to the melting temperature according to eq 3 (*28*).

$$[\text{urea}]_{1/2} = -\frac{1}{m}\left[\Delta H(T_{trs})\frac{T_{trs}-T}{T_{trs}} - \Delta C_p(T_{trs}-T) + \Delta C_p T \ln\left(\frac{T_{trs}}{T}\right)\right] \quad (3)$$

It has been shown by Myers et al. (*26*) that ΔC_p and the *m* value are linearly correlated at least by an empirical relationship. Consequently, the invariability of ΔC_p implies the invariability of *m* and vice versa.

Crystallization, X-ray Data Collection, and Structure Determination. Crystals of the various RNase T1 mutants in complex with the inhibitor 2′GMP were grown using macroseeding as described previously (*29*). Two crystal forms were obtained, one being the same as that reported for the wild-type 2′GMP complex (form 1) and the other having the same space group but a different unit cell (form 2). All data were collected at room temperature using a MAR Image Plate detector and a RIGAKU rotating anode X-ray generator operated at 40 kV and 90 mA, except for the V89S and V78A mutants, which were collected at EMBL beamline X11 of the DESY synchrotron (Hamburg, Germany). High-resolution data for the wild-type protein were collected at DESY beamline BW7B. The data were integrated using DENZO and subsequently scaled and merged using SCALEPACK (*30*). The CCP4 program TRUNCATE (*31*) was used to convert *I*'s into *F*'s. The statistics for the data collection are summarized in Table 1.

Form 1 crystal structures were determined by isomorphous substitution using the coordinates of a wild-type 2′GMP complex (PDB entry 5bu4) as a starting model. Form 2 crystal structures were determined by molecular replacement using the program AMORE (*32*) using the same starting structure. Refinement was carried out with X-PLOR version 3.851 (*33*) without any low-resolution or $I/(\sigma)I$ cutoff (or CNS version 1.0 in the case of the V16A mutant form 1 crystal). In each case, the refinement was started with a slow-cool stage to reduce bias toward the starting model. A bulk solvent correction was calculated after each rebuilding session. Water molecules were included (1) if they could be identified as peaks of at least 3σ in $F_o - F_c$ maps, (2) if they reappeared after refinement in $2F_o - F_c$ maps of at least 1σ, (3) if they made at least one reasonable hydrogen bond to a protein or inhibitor atom, and (4) if they showed no van der Waals clashes with any protein or inhibitor atom or with a previously identified water molecule. The quality of all refined structures was verified with PROCHECK (*34*). Coordinates have been deposited in the Protein Data Bank. PDB accession codes and refinement statistics are given in Table 1.

Crystal Structure Analysis. Sequence alignments of related RNases were performed manually, taking into account the conserved structural elements of the proteins. For structure superposition and visual analysis, the graphics program TURBO was used. Solvent accessibilities were calculated using the program NACCESS according to the method of Lee and Richards (*35*), using a probe sphere with a 1.4 Å radius. Hydrogen bonds were calculated using HBPLUS using D−A distances of 3.4 Å and the default minimum D−H−A, H−A−AA, and D−A−AA angles. The rms values of pairwise superpositions were calculated using the CCP4 program LSQKAB.

RESULTS

Stability Measurements

We constructed 12 single mutants in the hydrophobic core of RNase T1: Ala, Ser, Thr, and Cys at positions 16, 78, and 89 (Figure 1b). Of these, 11 mutants could be expressed at reasonable levels in *E. coli* via standard procedures (*19*). The relevant stability parameters obtained by urea unfolding and thermal unfolding followed by differential scanning calorimetry or circular dichroism are summarized in Table 2. Examples of the experimentally obtained urea denaturation curves and DSC thermograms are given in Figure 3. The Gibbs free energy of unfolding of the wild-type corresponds well with the results in other studies (*17*, *36*). All mutants are considerably less stable than the wild-type protein. When comparing the data to related work available in the literature, we found that the destabilizing effects of our mutations are at the upper range of what was measured for similar substitutions in other proteins. Destabilizations of similar magnitudes have until now always been associated with the recruitment of a novel buried water molecule (*15*). In our study, no water is buried in the hydrophobic core as a result of any of the current mutations in those cases where a crystal structure could be determined (see below).

DSC allows the experimental determination of the enthalpy effect and the calculation of the entropic parameter. Due to low expression levels, DSC experiments could only be carried out with the wild-type and five of the mutants. The data in Table 2 show large variations in enthalpy and entropy contributions, depending upon the site and nature of the mutation. A strong enthalpy−entropy compensation is also observed. These large and compensating enthalpy and entropy effects are characteristic of the hydrophobic effect (*37*, *38*). A general trend and the detailed physicochemical interpretation of these data cannot be obtained from measurements of only five mutants. Moreover, the low reversibility for the V16C mutant complicates a clear-cut interpretation of these data.

Urea Denaturation versus Thermal Denaturation

Figure 2a shows a plot of the urea half-point versus the melting temperature for the wild-type enzyme and the different mutants. A theoretical $[\text{urea}]_{1/2}$ versus T_{trs} plot according to eq 3 is superimposed. For most mutants, the thermodynamic parameters from urea denaturation curves, thermal denaturation curves, and DSC experiments are in good agreement, as has been shown before for other RNase T1 mutants (*10*, *17*). In the case of V16C, V89C, V78A, and V78T, the thermal and thermodynamic stability are clearly affected to a different degree. This can be explained by a temperature-dependent change in the heat capacity difference for those mutants. Such an explanation is also in agreement with the aberrant temperature dependence of the urea denaturation data of V16C, V89C, V78A, and V78T compared to the other mutations (Figure 2b−d). Changes in ΔC_p due to aggregation could be excluded experimentally; dynamic light scattering measurements yielded the molecular mass of the monomeric protein before unfolding and after refolding (data not shown).

Table 1: X-ray Data Collection Statistics

	unit cell dimensions (Å)											
	a	*b*	*c*	resolution (Å)	N_{meas}	N_{unique}	completeness (%)	R_{merge}	$I/\sigma(I)$	*R*	R_{free}	PDB code
wild type	40.32	47.61	50.54	40.0−1.23	151692	27375	94.6	0.063	16.52	0.178	0.192	1i0v
V16A (1)	40.23	46.59	49.76	26.0−1.80	70001	8704	93.1	0.101	18.58	0.190	0.209	1i2e
V16A (2)	40.39	42.77	56.66	15.0−1.95	29804	7382	97.6	0.138	8.42	0.172	0.205	1i2f
V16S	40.40	47.04	50.06	50.0−1.87	73803	7969	94.9	0.160	11.73	0.178	0.211	1g02
V16T	40.41	42.68	56.70	34.0−1.85	45840	8425	95.5	0.079	18.79	0.163	0.197	1i2g
V16C	40.44	42.68	56.65	20.0−2.00	26858	6623	94.3	0.127	9.22	0.169	0.208	1fys
V78A	40.43	42.72	56.78	15.0−1.80	36371	9514	99.1	0.099	11.99	0.167	0.198	1fzu
V78T	40.54	46.93	50.25	20.0−1.76	48824	9319	93.8	0.099	13.55	0.172	0.202	1i3i
V89S	40.32	47.04	50.13	20.0−2.35	13465	4204	98.6	0.151	7.11	0.173	0.283	1i3f

Table 2: Thermodynamic Data for Point Mutations at Positions 16, 78, and 89[a]

	T_{trs} (CD) (°C)	T_{trs} (DSC) (°C)	reversibility (DSC)	$\Delta\Delta G$ (25 °C, urea) (kcal/mol)	$\Delta\Delta G$ (25 °C, DSC) (kcal/mol)	$\Delta\Delta H$ (25 °C, DSC) (kcal/mol)	$T\Delta\Delta S$ (25 °C, DSC) (kcal/mol)
wild type	51.2	51.5	0.94	–	–	–	–
V16A	43.0	44.5	0.98	−2.49	−2.19	−5.57	−3.38
V16S	35.6	36.9	1.03	−4.66	−4.71	−25.77	−21.05
V16T	39.9	42.1	1.04	−3.20	−4.10	−30.37	−26.27
V16C	41.6	42.6	0.18	−5.30	−4.68	−42.17	−37.49
V78A	43.3	nd	nd	−4.08	nd	nd	nd
V78S	34.6	nd	nd	−4.73	nd	nd	nd
V78T	41.4	40.2	0.98	−4.08	−3.10	−3.52	−0.43
V78C	nd	nd	nd	−3.67	nd	nd	nd
V89S	29.6	nd	nd	−4.87	nd	nd	nd
V89T	38.0	nd	nd	−3.07	nd	nd	nd
V89C	46.3	nd	nd	−3.54	nd	nd	nd

[a] nd, not determined. Only for the wild type and five of the mutants could sufficient material be produced to allow DSC experiments. For the V78C mutant, the available material was even more limited, and thermal denaturation experiments using CD spectroscopy could also not be carried out.

Structural Consequences of the Mutations

For seven of the mutants (including all the position 16 mutants), a crystal structure could be obtained (Table 1). The overall structures of all seven of these mutants are very similar to that of wild-type RNase T1 (rms deviations for all atoms vary between 0.34 and 0.6 Å). The side chains of the mutants are accommodated very well with only minor local structural adaptations (Figures 4−6). The relevant structural parameters are given in Table 3. In contrast to the wild-type protein, residue 78 is present in a single conformation in most mutants. The only exception is V78T where Thr78 is seen in two conformations. This can be rationalized by considering that all other mutants at positions 16 and 89 for which a structure is available create extra cavity volume. The conformation of Val78 is in each case the one that results in a single larger cavity rather than two smaller ones. This minimizes the size of the cavity hydrophobic surface.

V16A and V78A. Both mutants can be considered pure cavity-creating mutants. Little or no adaptation of the immediately surrounding residues is observed that reduces the volume of the induced cavity (Figures 4a and 5a).

V16T and V78T. In the V16T mutant, the side chain of Thr16 has a rotamer conformation different from that of the original valine (Figure 4b). This allows for the Thr OH group to hydrogen bond with an ordered water molecule bridging Ser17 N and Ser14 O on the surface of the protein. This water molecule is also present in the wild-type structure and is thus not a consequence of the introduction of the Thr side chain. It is, however, packing-dependent as it is not found in other crystal forms of RNase T1 (*39*).

In contrast to V16T, the new threonine side chain of V78T cannot rearrange itself into a conformation where its hydroxyl group becomes solvent accessible. Like that in the wild-type structure, Thr78 is disordered and adopts two discrete rotamer conformations (Figure 5b). Only in one of these conformations is a potential hydrogen bond (oxygen−oxygen distance of 3.21 Å) introduced with the main chain oxygen of Trp59. Given the multiple conformations observed for Thr78, this hydrogen bond is not expected to contribute significantly to the stability of the V78T mutant.

V16S and V89S. The V16S mutant is completely isostructural with V16T. The introduced hydroxyl group is positioned in the same orientation as the threonine hydroxyl of V16T (Figure 4c). Consequently, V16S is also capable of hydrogen bonding to the same ordered water molecule at the surface of the protein. No additional compression of the structure is seen to relieve the stress of the enlarged internal hydrophobic cavity.

In V89S, the hydroxyl group of Ser89 is found in two orientations, corresponding to the orientations of both removed CH_3 groups (Figure 6). In one of these orientations, a potential hydrogen bond is formed with the π-electrons of Tyr4 (the distance between Ser89 OG and the center of the aromatic ring is 3.8 Å). In the other conformation, no hydrogen bond is introduced. We also observe a change in the orientation of the side chain of Val16 (Figure 6).

V16C. The SH group of Cys16 replaces the CG2 atom of the original valine (Figure 4d). The small cavity resulting from the removal of the CG1 atom is partly filled by a small (0.5 Å) movement of Leu86. Apart from this, no significant differences between the mutant and wild-type protein are seen.

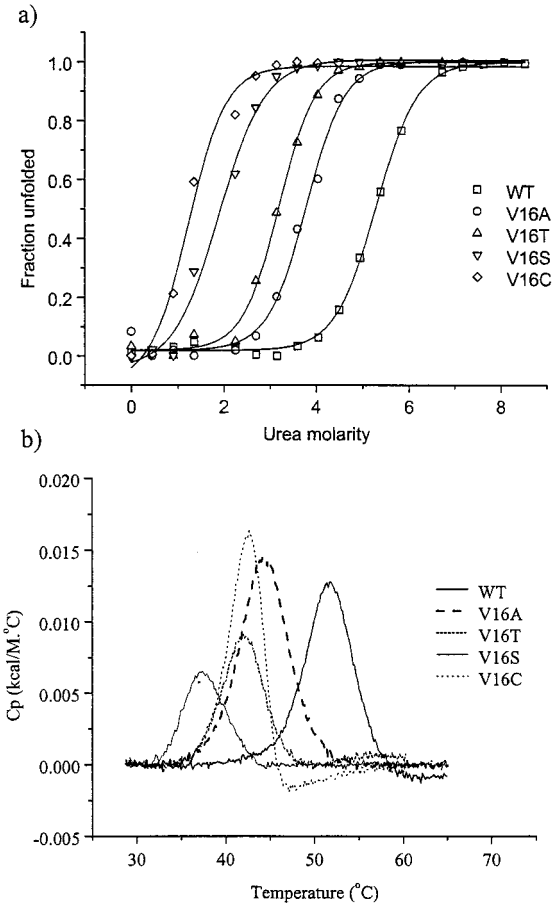

FIGURE 3: Examples of experimental unfolding data. (a) Urea unfolding curves of wild-type RNase T1, V16A, V16S, V16T, and V16C. (b) DSC thermograms for the same proteins.

DISCUSSION

Structural Context of Val16, Val78, and Val89 in Wild-Type RNase T1. The three residues selected in this study (Val16, Val78, and Val89) are completely buried in the major hydrophobic core of RNase T1 and make extensive contacts with each other (Figure 1b). Val78 is located approximately at the center of mass of the enzyme. Exposing the side chain of Val78 to solvent probably requires an almost complete unfolding of the enzyme. Val16 and Val89 are also completely buried, but their position closer to the protein surface suggests that their side chains can be at least partially exposed to solvent during and via local structural rearrangements.

Val78 is one of the best conserved residues outside the active site within this family of microbial ribonucleases (*39*, *40*). Even in the distantly related barnase and RNase Sa, this Val residue is substituted only with Leu and Ile, respectively. Val16 and Val89 are only conserved within the fungal ribonucleases that are closely related to RNase T1 and have no true structural counterpart in barnase and RNase Sa.

Secondary Structure Propensity, Packing Density, and Hydrophobic Effect. Val16 and Val89 are located in the α-helix and a loop segment, respectively, while Val78 is anchored in the central strand of the *β*-sheet pointing directly toward the body of the helix. As a consequence, secondary structure propensities may contribute to the observed differences in the conformational stabilities of the mutants. Knowing that the hydrophobic core of RNase T1 is efficiently packed, one would expect tertiary or packing interactions to play a crucial role in the stability in this part of the protein, as well (*41*−*43*). When the ensemble of available data was analyzed, no correlations could be found between the intrinsic secondary structure propensities of the different amino acids (*44*−*46*) and the mutational effects on the stability parameters. Also, the measured differences in stability relative to the wild-type enzyme are much larger than the values expected from secondary structure propensi-

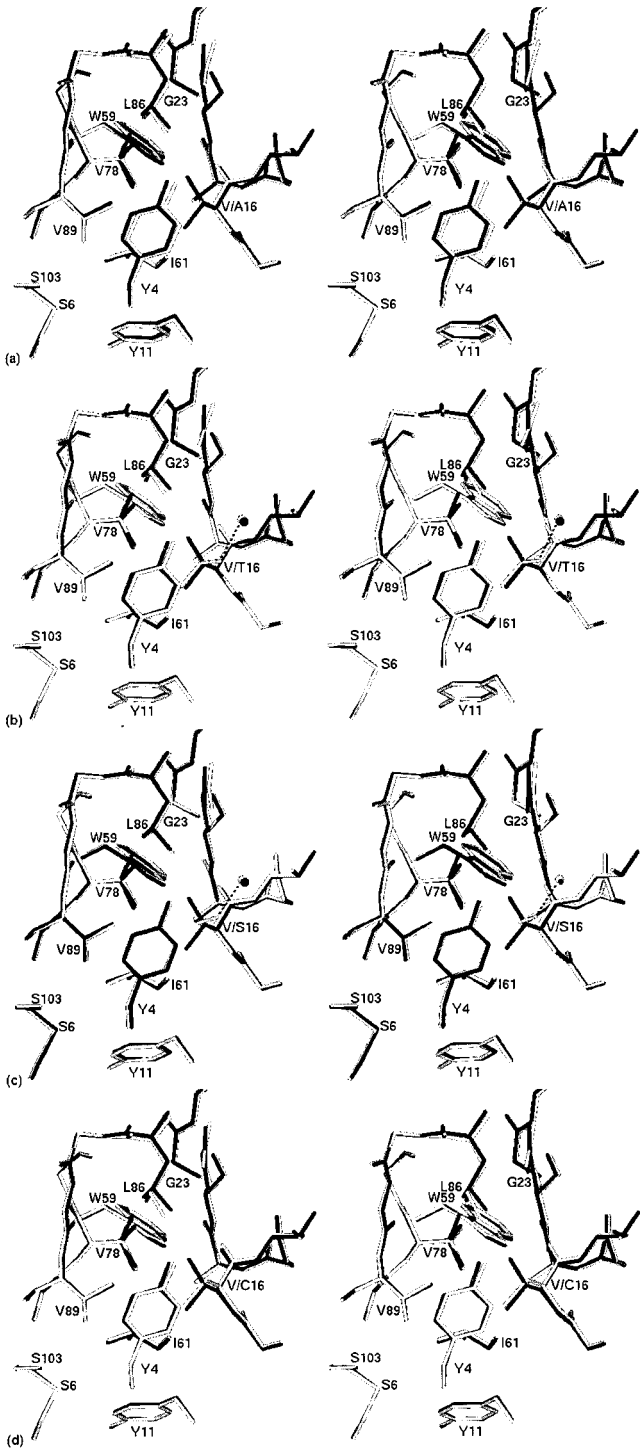

FIGURE 4: Structural consequences of mutating Val16. In each instance, a superposition is shown of the mutant structure (light ball-and-stick representation) and the wild-type protein (dark ball-and-stick representation): (a) V16A mutation, (b) V16S mutation, (c) V16T mutation, and (d) V16C mutation.

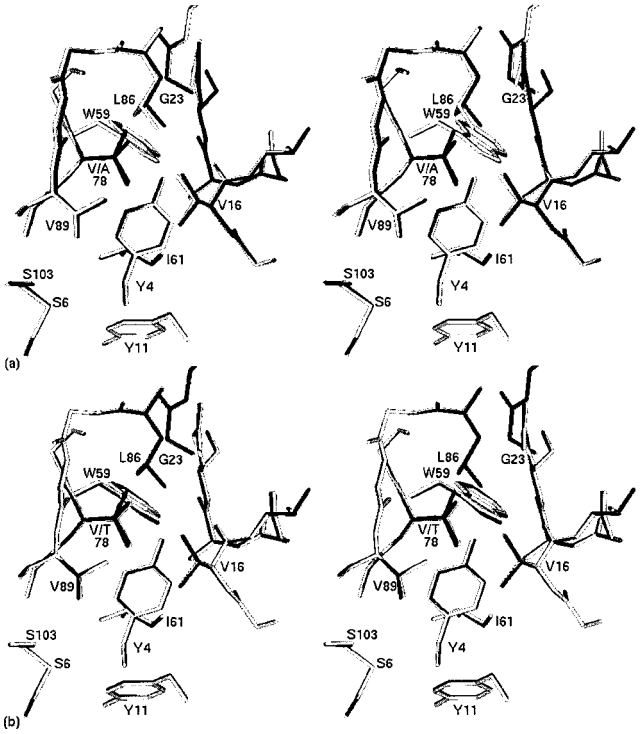

FIGURE 5: Structural consequences of mutating Val78. In each instance, a superposition is shown of the mutant structure (light ball-and-stick representation) and the wild-type protein (dark ball-and-stick representation): (a) V78A mutation and (b) V78T mutation.

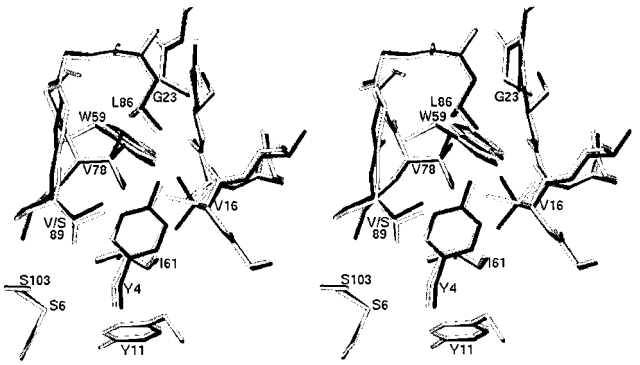

FIGURE 6: Structural consequences of mutating Val89 to Ser. A superposition is shown of the mutant structure (light ball-and-stick representation) and the wild-type protein (dark ball-and-stick representation).

ties alone (*45*−*47*). Thus, each substitution of a highly buried valine residue seems to result in a reduced stability primarily due to the hydrophobic effect and to a lesser extent to the loss of favorable van der Waals interactions. This is in agreement with the observations of other groups (*4*, *48*).

Energetics of Burying CH_3, OH, and SH Groups in Small Pre-Existing Hydrophobic Cavities. The highly conservative nature of the mutations as far as three-dimensional structure is concerned allows us to calculate the energetic gain or cost for the introduction of a CH_3, OH, or SH group into a small pre-existing hydrophobic cavity. The energetic consequences of introducing a methyl group can be analyzed through Ala to Val or Ser to Thr substitutions providing the gain in stability of introducing one or two CH_3 groups into a pre-existing cavity. Similarly, the impact of adding a hydroxyl or sulfhydryl group to a preformed cavity can be investigated via an Ala to Ser or Ala to Cys mutation, respectively. The results from such calculations are shown in Table 4. Two important observations become apparent.

Table 3: Structural Consequences of Mutations at Positions 16, 78, and 89

	H-bonds introduced	rms with wild type (Å) (all/backbone/ 5 Å sphere)	solvent accessibility of the relevant side chains (Å^2) (16/78/89)
wild type	–	–	2.74/0.00/0.00
V16A (1)	none	0.459/0.185/0.224	0.08/0.00/0.00
V16A (2)	none	0.458/0.303/0.375	1.91/0.00/0.40
(1) vs (2)		0.604/0.305/0.102	
V16T	Thr16 OH···Wat	0.427/0.300/0.167	4.00/0.00/0.38
V16S	Ser16 OH···Wat	0.429/0.166/0.192	2.57/0.00/0.00
V16C	none	0.437/0.311/0.183	3.50/0.00/0.26
V78A	none	0.403/0.279/0.146	3.42/0.00/0.32
V78T	Thr78 OH···Trp59 O	0.337/0.148/0.095	2.89/0.00/0.00
V89S	Ser89 OH···Tyr4π or Ser89 OH···Cys6 SG	0.450/0.224/0.260	3.29/0.00/0.00

First, while the introduction of a CH_3 or OH group has roughly the same consequences at all three positions, the effect of cysteine is much more variable. Filling a hydrophobic cavity with a SH group can have a stabilizing as well as a destabilizing effect. This is rather unexpected for two reasons. First, the microenvironments around Val16, Val78, and Val89 are quite similar. Second, in hydrophobicity scales derived from a large variety of experiments, Cys consistently behaves as a hydrophobic amino acid, and is also considered as such in the literature (*49*). The destabilizing effect is especially pronounced in the V16C mutation, which is the most destabilizing of all. However, when the crystal structures of the wild-type, V16C, and V16S are compared, the structures of the V16C and V16S mutants clearly show that the side chain of residue 16 has the possibility of reorienting itself to make a hydrogen bond with a solvent water molecule. This is observed in the V16S mutant in agreement with the hydrophilic nature of an OH group, but not in the V16C mutant. The (weak) hydrogen bond potential of Cys, which is not fulfilled in the V16C mutation, is thus not the prime reason for the unexpectedly low stability of this mutant.

The second major observation is that the penalty for introducing an OH group into a particular hydrophobic cavity is roughly the same as the gain obtained from filling the cavity with a CH_3 group. Matthews and co-workers previously observed that, in general, the magnitude of the destabilization of Val → Thr mutations in T4 lysozyme is similar to that of Val → Ala mutations (*15*), but they did not systematically create Val → Thr and Val → Ala mutations at the same position. When looking in the literature for such cases where the thermodynamic stability was measured for Val → Thr as well Val → Ala mutations (or Thr → Val and Thr → Ala) at identical sites, we found only 20 cases (including those from the study presented here). Figure 7 plots the unfolding $\Delta\Delta G$ of the 13 Val → Thr versus Val → Ala mutations as well as the seven Thr → Val versus Thr → Ala mutations. These experimental data points are close to the diagonal of the plot, as is expected for an inverse behavior of a CH_3 and an OH group for filling a pre-existing cavity (correlation = 0.88; slope = 0.87). This is despite the diverse structural environment of the mutated sites (from completely surrounded by hydrophobics over more polar

Table 4: Energetics of Burying CH_3, OH, and SH Groups in a Small Pre-Existing Hydrophobic Cavity[a]

		energetic cost (kcal/mol)		
mutation	addition	position 16	position 78	position 89
Ala → Val	two CH_3 groups	−1.25 kcal/CH_3	−2.0 kcal/CH_3	−
Ser → Thr	CH_3 group	−1.5 kcal/CH_3	−1.5 kcal/CH_3	−1.15 kcal/CH_3
Ala → Ser	OH group	>2.2 kcal/OH	1.3 kcal/OH	2.2 kcal/OH[b]
Ala → Cys	SH group	2.5 kcal/SH	−0.4 kcal/SH	1.5 kcal/SH[b]

[a] The values are derived from the urea denaturation experiments and are averages from the data collected at five different temperatures. [b] Assuming the Val → Ala mutation costs ~2.3 kcal/mol as determined from the Thr → Ser mutation.

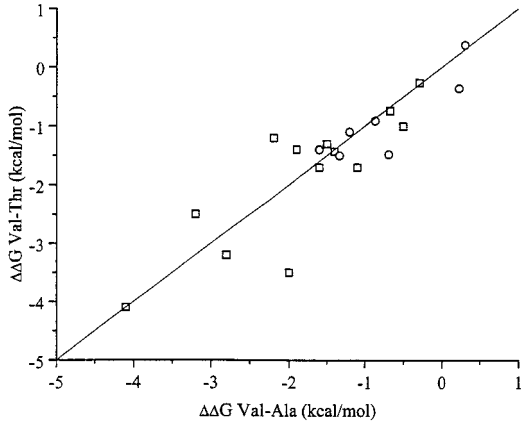

FIGURE 7: Comparison of the destabilization of Val → Thr with Val → Ala (□) and Thr → Val with Ala → Val (○). The values that are shown are from positions 16 and 78 of RNase T1 (this study), positions 3, 34, 36, 60, and 63 of chymotrypsin (*72*), positions 10, 36, and 45 of barnase (*73*), positions 87 and 149 of T4 lysozyme (*9*, *74*), and positions 11, 40, 43, 52, 70, 93, 99, and 100 of human lysozyme (*60*, *75*). The straight line that is shown is not a least-squares fit but simply the diagonal of the plot, which would be the expected prediction if the penalty of introducing an OH group into a particular hydrophobic cavity is roughly the same as the gain obtained from filling the cavity with a CH_3 group or vice versa.

toward highly solvent exposed), the differences in experimental conditions, and their origin in different proteins.

It is our opinion that this inverse equivalence of the behavior of CH_3 and OH groups is crucial for the existence of proteins with a stable and unique three-dimensional fold. The inverse equivalence causes the burial of a hydrophobic residue to be favorable and the burial of a hydrophilic residue to be penalized. This is consistent with the concept of negative design elements.

A fold not only is stabilized by a number of stabilizing interactions but also must contain a large number of negative design elements. Initial studies aiming at de novo design resulted in polypeptides that contained a great deal of the desired secondary structure contents, but had no defined tertiary structure (*50*−*52*). Indeed, these earlier attempts focused on elements that stabilize particular (sub)structures, ignoring the need to also prevent other undesired folds from becoming populated. For a stable folded protein, not only the free energy difference between the folded and the unfolded state but also that between the correct fold and (partially) misfolded structures must be sufficiently large (*53*). Modern design exercises taking into consideration not only positive but also negative design elements are much more successful. The first study to incorporate these ideas involved the design of a four-helix bundle with left-turning geometry (*53*). Since then, several truly successful designs have been presented, including designs meeting the Paracelsus challenge: changing from one stable fold to another by changing less than 50% of the amino acids of a known protein (*54*−*56*).

Negative design must also be present in naturally occurring proteins. Recent work on the Arc repressor has shown that the balance between positive and negative design elements is a delicate one (*57*). A single hydrophilic to hydrophobic mutation is sufficient to create a protein that switches between two folds. A second hydrophobic to hydrophilic mutation can stabilize the new fold compared to the native one.

It is therefore of interest to learn how nature solves the problem of combined negative and positive design. An attractive hypothesis is that this is done in a universal way using a property present in all proteins. Mirror-like energetics of OH versus CH_3 burying may be such a universal tool. If the energetics of OH burying compared to CH_3 burying would be significantly different in magnitude, either the interactions in the hydrophobic core or the interactions of the surface residues with the surrounding solvent would become a dominating factor, making use of negative design elements to select a unique fold much more difficult.

Comparison with Related Studies. Yutani and co-workers developed an empirical model for predicting the thermodynamic stability of point mutants of human lysozyme (*58*):

$$\Delta\Delta G_{est} = \Delta\Delta G_{HP} + \Delta\Delta G_{conf} + \Delta\Delta G_{HB} + \Delta\Delta G_{H_2O} \quad (4)$$

where $\Delta\Delta G_{HP} = \alpha\Delta\Delta ASA_{NP} + \beta\Delta\Delta ASA_P$, $\Delta\Delta G_{HB} = \gamma\Sigma r_{HB} - 1$, $\Delta\Delta G_{H_2O} = \delta\Delta N_{H_2O}$.

We therefore calculated in a similar way $\Delta\Delta G_{est}$ for the seven RNase T1 mutants for which a crystal structure is available. Interestingly, the calculated values for the conformational stability ($\Delta\Delta G_{est}$) for the RNase T1 mutants do not predict the experimentally observed values ($\Delta\Delta G_{exp}$) (data not shown). For all but two of the mutants (V78T and V89S), the model even predicts an increase in stability. Test calculations on human lysozyme mutants on the other hand, for which the model was developed (*58*), showed a good agreement between experimental and calculated stabilities.

Yutani and co-workers also applied their model to the vast amount of mutational data available in the literature for T4 lysozyme (*59*). The parameters of the various stabilization factors derived from mutant human lysozymes are fairly compatible with the mutant T4 lysozymes. An optimal correlation, however, required reparametrization of eq 4 as well as the use of additional terms. Given the fact that the current study includes only seven RNase T1 mutants with known structure, development of an RNase T1 specific model is not feasible.

Lately, studies have been published that challenge the dominating role of the hydrophobic effect as the driving force for protein folding. In a mutational study of human lysozyme, Takano and co-workers concluded that polar group burying has only a low energetic penalty, and that hydrogen bonding is the major determinant of protein stability (*60*−*62*). Similarly, from studies of Asn to Ala, Leu, and Ile mutations (*63*−*65*), Pace concluded that "the burial of an amide group makes a larger contribution to protein stability than an equivalent volume of -CH_2- groups" (*66*).

The major difference between our study and the other studies in which hydrophobic to polar (and vice versa) substitutions were investigated is the microenvironment of the side chains that were considered. Our study is the only one of which we know of where the burial of OH, SH, and CH_3 groups was assessed in a true hydrophobic environment, i.e., one where no new hydrogen bonds can be formed with the introduced OH group. It can be expected that the energetics of substituting a small hydrophobic residue with a hydrophilic one will depend on the microenvironment in which this substitution occurs. The mutations in the studies by Takano and Pace all introduce good hydrogen bonds. This results in a significantly reduced penalty. This might also contribute to why the calculated stability values ($\Delta\Delta G_{est}$) based on human lysozyme-derived parameters do not predict the experimental stabilities for our RNase T1 mutants.

A large fraction of the penalty of burying hydrophilic groups thus comes from shielding them from potential hydrogen bonding partners. A large fraction of hydrophilic side chains and peptide groups are buried in the interior of proteins (*67*, *68*). They invariantly are involved in hydrogen bonds. The question one needs to ask is therefore how large the contribution of these buried hydrogen bonds to protein stability is. Most studies measure the strength of a hydrogen bond from unfolding data from a single mutation. These may provide quite large differences in unfolding free energy, but cannot measure the contribution of a single hydrogen bond. The correct way to measure the strength of an individual interaction is through a double-mutant cycle (*69*). Isolated buried hydrogen bonds contribute little if anything to stability if measured via double-mutant cycles, but can contribute up to 1−1.5 kcal/mol when involved in a larger network formed by a cluster of buried hydrophilic side chains (*70*). When these values are compared with those for the destabilization of the Val → Thr and Val → Ser mutations in the study presented here, it seems that the hydrophobic effect, the partitioning of hydrophilic and hydrophobic groups, is a very important factor in protein stability, a conclusion also reached in a recent related study by Matthews and co-workers (*71*).

ACKNOWLEDGMENT

We acknowledge the use of synchrotron beamtime at EMBL beamlines X11 and BW7B at the DORIS storage ring (Hamburg, Germany). We thank Elke Brosens for technical assistance and Lieven Buts for helpful discussions and critical reading of the manuscript.

REFERENCES

1. Dill, K. A. (1999) *Protein Sci. 8*, 1166−1180.
2. Anfinsen, C. B. (1973) *Science 181*, 223−230.
3. Knowles, J. R. (1987) *Science 236*, 1252−1258.
4. Buckle, A. M., Cramer, P., and Fersht, A. R. (1996) *Biochemistry 35*, 4298−4305.
5. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) *Science 255*, 178−183.
6. Wilcock, D., Pisabarro, M. T., Lopez-Hernandez, E., Serrano, L., and Coll, M. (1998) *Acta Crystallogr. D54*, 378−385.
7. Takano, K., Yamagata, Y., Fujii, S., and Yutani, K. (1997) *Biochemistry 36*, 688−698.
8. Byrne, M. P., Manuel, R. L., Lowe, L. G., and Stites, W. E. (1995) *Biochemistry 34*, 13949−13960.
9. Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W., and Matthews, B. W. (1993) *Biochemistry 32*, 11363−11373.
10. Shirley, B. A., Stanssens, P., Hahn, U., and Pace, C. N. (1992) *Biochemistry 31*, 725−732.
11. Serrano, L., Kellis, J. T. J., Cann, P., Matouschek, A., and Fersht, A. R. (1992) *J. Mol. Biol. 224*, 783−804.
12. Lim, W. A., Farruggio, D. C., and Sauer, R. T. (1992) *Biochemistry 31*, 4324−4333.
13. Kauzmann, W. (1959) *Adv. Protein Chem. 14*, 1−63.
14. Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
15. Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W., and Matthews, B. W. (1993) *Biochemistry 32*, 11363−11373.
16. Steyaert, J. (1997) *Eur. J. Biochem. 247*, 1−11.
17. Langhorst, U., Backmann, J., Loris, R., and Steyaert, J. (2000) *Biochemistry 39*, 6586−6593.
18. Pace, C. N., Heinemann, U., Hahn, U., and Saenger, W. (1991) *Angew. Chem., Int. Ed. 30*, 343−454.
19. Steyaert, J., Hallenga, K., Wyns, L., and Stanssens, P. (1990) *Biochemistry 29*, 9064−9072.
20. Chen, B., and Przybyla, A. E. (1994) *BioTechniques 17*, 657−659.
21. Mayr, L. M., and Schmid, F. X. (1993) *Protein Expression Purif. 4*, 52−58.
22. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
23. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
24. Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
25. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
26. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
27. Privalov, P. L., and Khechinashvili, N. N. (1974) *J. Mol. Biol. 86*, 665−684.
28. Backmann, J., Schäfer, G., Wyns, L., and Bönisch, H. (1998) *J. Mol. Biol. 284*, 817−833.
29. Zegers, I., Loris, R., Dehollander, G., Fattah, H. A., Poortmans, F., Steyaert, J., and Wyns, L. (1998) *Nat. Struct. Biol. 5*, 280−283.
30. Otwinowski, Z., and Minor, W. (1997) in *Methods in Enzymology* (Carter, C. W., and Sweet, R. M., Eds.) pp 307−326, Academic Press, New York.
31. Collaborative Computational Project No. 4 (1994) *Acta Crystallogr. D50*, 760−763.
32. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
33. Brünger, A. T. (1992) *X-PLOR: A system for crystallography and NMR*, version 3.1, Yale University Press, New Haven, CT.
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
35. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.
36. Pace, C. N., Shirley, B. A., and Thomson, J. A. (1989) in *Protein structure: A practical approach* (Creighton, T. E., Ed.) pp 311−330, Oxford University Press, New York.
37. Creighton, T. E. (1993) in *Proteins: Structure and Molecular Properties*, W. H. Freeman and Company, New York.
38. Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
39. Loris, R., Langhorst, U., De Vos, S., Decanniere, K., Bouckaert, J., Maes, D., Transue, T. R., and Steyaert, J. (1999) *Proteins 36*, 117−134.

40. Hill, C., Dodson, G., Heinemann, U., Saenger, W., Mitsui, Y., Nakamura, K., Borisov, S., Tischenko, G., Polyakov, K., and Pavlovsky, S. (1983) *Trends Biochem. Sci. 8*, 364−369.
41. Ratnaparkhi, G. S., and Varadarajan, R. (2000) *Biochemistry 39*, 12365−12374.
42. Vlassi, M., Cesareni, G., and Kokkinidis, M. (1999) *J. Mol. Biol. 285*, 817−827.
43. Sneddon, S. F., and Tobias, D. J. (1992) *Biochemistry 31*, 2842−2846.
44. Koehl, P., and Levitt, M. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 12524−12529.
45. Blaber, M., Zhang, X. J., and Matthews, B. W. (1993) *Science 260*, 1637−1640.
46. Smith, C. K., Withka, J. M., and Regan, L. (1994) *Biochemistry 33*, 5510−5517.
47. Kim, C. A., and Berg, J. M. (1993) *Nature 362*, 267−270.
48. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) *Science 255*, 178−183.
49. Jesior, J. C. (2000) *J. Protein Chem. 19*, 93−103.
50. Goraj, K., Renard, A., and Martial, J. A. (1990) *Protein Eng. 3*, 259−266.
51. Tanaka, T., Kimura, H., Hayashi, M., Fujiyoshi, Y., Fukuhara, K., and Nakamura, H. (1994) *Protein Sci. 3*, 419−427.
52. Ilyina, E., Roongta, V., and Mayo, K. H. (1997) *Biochemistry 36*, 5245−5250.
53. Hecht, M. H., Richardson, J. S., Richardson, D. C., and Ogden, R. C. (1990) *Science 249*, 884−891.
54. Jones, D. T., Moody, C. M., Uppenbrink, J., Viles, J. H., Doyle, P. M., Harris, C. J., Pearl, L. H., Sadler, P. J., and Thornton, J. M. (1996) *Proteins 24*, 502−513.
55. Yuan, S. M., and Clarke, N. D. (1998) *Proteins 30*, 136−143.
56. Dalal, S., Balasubramanian, S., and Regan, L. (1997) *Nat. Struct. Biol. 4*, 548−552.
57. Cordes, M. H., Burton, R. E., Walsh, N. P., McKnight, C. J., and Sauer, R. T. (2000) *Nat. Struct. Biol. 7*, 1129−1132.
58. Funahashi, J., Takano, K., Yamagata, Y., and Yutani, K. (1999) *Protein Eng. 12*, 841−856.
59. Funahashi, J., Takano, K., and Yutani, K. (2001) *Protein Eng. 14*, 127−134.
60. Takano, K., Yamagata, Y., and Yutani, K. (2001) *Biochemistry 40*, 4853−4858.
61. Funahashi, J., Takano, K., Yamagata, Y., and Yutani, K. (2000) *Biochemistry 39*, 14448−14456.
62. Takano, K., Tsuchimori, K., Yamagata, Y., and Yutani, K. (2000) *Biochemistry 39*, 12375−12381.
63. Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) *FASEB J. 10*, 75−83.
64. Pace, C. N., Hebert, E. J., Shaw, K. L., Schell, D., Both, V., Krajcikova, D., Sevcik, J., Wilson, K. S., Dauter, Z., Hartley, R. W., and Grimsley, G. R. (1998) *J. Mol. Biol. 279*, 271−286.
65. Pace, C. N. (1995) *Methods Enzymol. 259*, 538−554.
66. Pace, C. N. (2001) *Biochemistry 40*, 310−313.
67. Lesser, G. J., and Rose, G. D. (1990) *Proteins 8*, 6−13.
68. Kajander, T., Kahn, P. C., Passila, S. H., Cohen, D. C., Lethiö, L., Adolfsen, W., Warwicker, J., Schell, U., and Goldman, A. (2000) *Structure 8*, 1203−1214.
69. Horovitz, A., and Fersht, A. R. (1990) *J. Mol. Biol. 214*, 613−617.
70. Albeck, S., Unger, R., and Schreiber, G. (2000) *J. Mol. Biol. 298*, 503−520.
71. Xu, J., Baase, W. A., Quillin, M. L., Baldwin, E. P., and Matthews, B. W. (2001) *Protein 10*, 1067−1078.
72. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) *J. Mol. Biol. 254*, 260−288.
73. Serrano, L., Kellis, J. T. J., Cann, P., Matouschek, A., and Fersht, A. R. (1992) *J. Mol. Biol. 224*, 783−804.
74. Xu, J., Baase, W. A., Baldwin, E., and Matthews, B. W. (1998) *Protein Sci. 7*, 158−177.
75. Takano, K., Yamagata, Y., Funahashi, J., Hioki, Y., Kuramitsu, S., and Yutani, K. (1999) *Biochemistry 38*, 12698−12708.

BI010565N