The Role of Packing Interactions in Stabilizing Folded Proteins

Scott F. Sneddon*,‡ and Douglas J. Tobias§

Central Research, Pfizer Inc., Groton, Connecticut 06379, and Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received August 8, 1991; Revised Manuscript Received October 17, 1991

ABSTRACT: In order to investigate the role of nonpolar side chains in determining protein stability, we have carried out a molecular dynamics simulation study of the thermodynamics of interconverting isoleucine and valine side chains in the core of ribonuclease T_1. The free energy change in the unfolded state, which we take to be fully solvated, was small and agrees qualitatively with experimental studies of alkane solvation. In the two Ile → Val mutations studied, the protein was able to relax around the smaller side chains, while in the case of the two Val → Ile mutations, the ability of the core to accommodate the extra methylene group depended on where the mutation took place. We argue that the experimentally observed decrease in stability for mutating isoleucine into valine results from a loss of favorable packing interactions of the side chain in the folded form of the protein. This supports the view that packing interactions in the folded state are an important contributor to the overall stability of the folded protein and that the core of the native protein is packed efficiently and almost completely.

A striking structural feature of globular proteins is that, for the most part, the nonpolar side chains are on the inside, sequestered away from the aqueous solvent, forming a dense hydrophobic "core". This observation has lead to much speculation regarding the role of core formation in stabilizing the folded protein. The explanation most often given, first proposed by Kauzmann (1959) before the first protein crystal structures were known, is that the aversion of nonpolar groups for water is the most important physical force stabilizing the folded state [for a review, see Dill (1990)]. Kauzmann's idea was that the formation of one "hydrophobic bond" between two nonpolar groups is accompanied by the formation of one full hydrogen bond between water molecules. Summed over the many nonpolar residues folded into the core, this process would result in a significant stabilization of the folded state. The evidence most often cited in support of this idea is the similarity between the thermodynamics of protein unfolding and the transfer of nonpolar substances from nonpolar solvents to water (Tanford, 1980; Dill, 1990). Another possibility is that the burial of nonpolar side chains in the interior of the protein is favorable not simply because of the aversion of nonpolar groups toward water but also because these groups can participate in favorable "packing interactions" with each other in the core of the folded protein (Richards, 1977; Kellis et al., 1988, 1989).[1] Indeed, on the basis of a novel thermodynamic analysis of protein unfolding and nonpolar transfer data, Murphy et al. (1990) have recently proposed that the stability of the folded state arises from favorable protein–protein interactions rather than from unfavorable protein–water interactions.

In order to learn more about the role of nonpolar residues in determining protein stability, several mutation studies have recently been carried out (Kellis et al., 1988, 1989; Lim & Sauer, 1989; Shortle et al., 1990; Sandberg & Terwilliger, 1991). Among other things, the results of these studies show that the removal of a single methylene group from the nonpolar core (e.g., from the mutation of isoleucine to valine) can result in a loss of stability of the folded state of between 1 and 2.5 kcal/mol (Kellis et al., 1988, 1989; Shortle et al., 1990; Sandberg & Terwilliger, 1991). This change in stability can be explained by a decrease in the aversion of the nonpolar group for water in the unfolded state (i.e., the unfolded state free energy decreases) and/or by the loss of favorable packing interactions in the folded state (i.e., the folded state free energy increases). The interpretation that is adopted will dramatically influence our understanding of protein structural stability and is still a matter of debate [see, for example, the varying interpretations of Richards (1977), Tanford (1980), Jaenicke (1987), Kellis et al. (1988, 1989), Dill (1990), Murphy et al. (1990), Pace (1990), Shortle et al. (1990), Behe et al. (1991), Gellman (1991), and Sandberg and Terwilliger (1991)].

The thermodynamic stability of the folded state, ΔG^f, is measured experimentally as the free energy difference between the folded and unfolded states (i.e., $\Delta G^f = G^{folded} - G^{unfolded}$). The change in stability caused by mutation, $\Delta\Delta G_{native \rightarrow mutant}$, is the difference in stability of the native and mutant proteins ($\Delta\Delta G_{native \rightarrow mutant} = \Delta G^f_{mutant} - \Delta G^f_{native}$). It is difficult to determine experimentally whether the observed $\Delta\Delta G$ for a mutation arises primarily from a change in the free energy of the folded state, G^{folded}, or the unfolded state, $G^{unfolded}$, or both. The free energy simulation method computes $\Delta\Delta G_{native \rightarrow mutant}$ as the free energy difference for mutation in the folded state minus the free energy for mutation in the unfolded state, i.e., as $\Delta G^{folded}_{native \rightarrow mutant} - \Delta G^{unfolded}_{native \rightarrow mutant}$ (Wong & McCammon, 1986; Dang et al., 1989; Hermans et al., 1990; Tidor & Karplus, 1991). Thus, the simulations provide estimates of the contributions from both the folded and unfolded states in changing the stability of the mutant protein.

In this paper we describe a study which employs free energy stimulation methods to suggest whether the decrease in stability caused by Ile → Val mutants in the protein interior arises from changes in the free energy of the folded or unfolded state. We have investigated both Ile → Val and Val → Ile mutations in the core of RNase T_1 and have characterized the structural changes which take place as a result of mutation, as well as the interactions which differ in the native and mutant proteins.

* Author to whom correspondence should be addressed.
‡ Pfizer, Inc.
§ University of Pennsylvania.

[1] Here we use the term "packing interactions" to refer to the favorable van der Waals interactions between nonpolar groups.

0006-2960/92/0431-2842$03.00/0

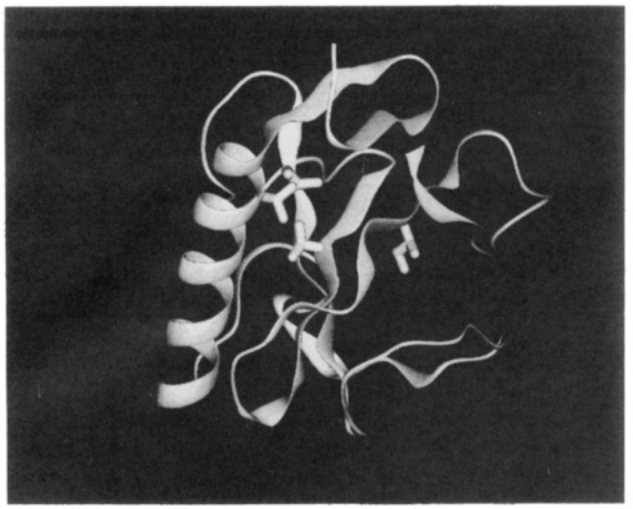

FIGURE 1: Ribbon diagram of the backbone structure of RNase T_1. The secondary structures are shown as flat ribbons, while the side chains that have been mutated are shown as thick tubes (all figures were produced with the QUANTA molecular graphics program, Polygen, Inc., Waltham, MA). The side chains of Val-16, Ile-61, and Val-78 form a common core, while Ile-90 projects into a smaller core on the opposite side of the β-sheet.

Our work offers an explanation in accord with those of Kellis et al. (1988, 1989) that the stability change from the Ile → Val mutation arises predominantly from changes in the free energy of the folded state.

Materials and Methods

RNase T_1 has several features which make it attractive for our study: it is relatively small (104 amino acids), it possesses a common α–β folding motif (Richardson, 1981), it is genetically related to the barnase enzyme that Kellis et al. (1988, 1989; Hill et al., 1983) have used in their mutation studies, and a high-resolution X-ray structure exists in the Brookhaven Protein Data Bank (resolution = 1.9 Å, R = 0.18; Arni et al., 1988). The mutations Ile → Val and Val → Ile were chosen because these side chains differ only by a single methylene group and because there is a wealth of experimental data for Ile → Val and Val → Ile mutations with which we can compare our results (Kellis et al., 1988, 1989; Shortle et al., 1990; Sandberg & Terwilliger, 1991). We have mutated Ile-61 and Ile-90 to valine and Val-16 and Val-78 to isoleucine to probe the influence of removing and adding methylene groups to the nonpolar core. The side chains of Ile-61, Val-16, and Val-78 all reside in the major core formed at the α–β interface, while Ile-90 is in a minor core on the opposite side of the β-sheet (Figure 1).

We have chosen as our model of the unfolded state a completely solvent-exposed blocked isoleucine residue (CH_3CO–Ile–$NHCH_3$). We consider the computed free energy change a limiting value for the unfolded state. If the true unfolded state more closely resembles the folded state (if the side chain is partially buried), then the change in overall stability for mutation will be smaller than calculated here. Nevertheless, we feel this model is a reasonable choice without more detailed information about the side-chain environment in the unfolded state.

We used the free energy simulation method described by Fleischman and Brooks (1987) to calculate $\Delta G^{\text{folded}}_{\text{native}\rightarrow\text{mutant}}$ and $\Delta G^{\text{unfolded}}_{\text{native}\rightarrow\text{mutant}}$ for Ile → Val and Val → Ile mutations. In our application of this method, a native side chain is gradually transformed into a mutant side chain in a series of five molecular dynamics simulations ("windows"). We carried out additional simulations of the native and mutant protein "end points" to characterize the structural and energetic effects of the mutations.

Table I: Free Energy Changes in the Unfolded and Folded States and Overall Stability Changes for Ile → Val and Val → Ile Mutations in RNase T_1[a]

mutation	$\Delta G^{\text{unfolded}}_{\text{native}\rightarrow\text{mutant}}$	$\Delta G^{\text{folded}}_{\text{native}\rightarrow\text{mutant}}$	$\Delta\Delta G_{\text{native}\rightarrow\text{mutant}}$
I61V	0.4 ± 0.1	2.3 ± 0.3	1.9 ± 0.3
I90V	0.4 ± 0.1	1.6 ± 0.1	1.2 ± 0.1
V16I	−0.4 ± 0.1	0.2 ± 0.1	0.6 ± 0.1
V78I	−0.4 ± 0.1	4.5 ± 0.3	4.9 ± 0.3

[a] Free energy changes are given in kcal/mol. Reported uncertainties are error-propagated standard deviations of 120 averages of 100 data sets (Fleischman & Brooks, 1987).

Each simulation consisted of 6000 steps equilibration and 12 000 steps of data collection. The Verlet (1967) algorithm was used to integrate Newton's equations of motion with a time step of 1.5 fs. The folded state simulations were carried out in vacuo starting from the X-ray coordinates of Arni et al. (1988). The unfolded state simulations were performed in a cubic box (edge length 18.856 Å) containing 198 water molecules with periodic boundary conditions. We used the CHARMM protein parameters (Brooks et al., 1983) and the three-site TIP3P model of Jorgensen et al. (1983) for water. The nonbonded energies and forces were smoothly truncated at 7.75 Å using a van der Waals switching function and an electrostatic shifting function (Brooks et al., 1983), based on atomic centers. The nonbonded interactions were processed using a list-based algorithm (Verlet, 1967), and the lists were updated every 10 steps. The velocities were periodically reassigned from a Maxwell–Boltzmann distribution to maintain temperatures of approximately 25 °C. The SHAKE constraint algorithm (Ryckaert et al., 1977) was used to keep the water molecules rigid and to maintain rigid N–H and O–H bonds in the protein. All of the simulations were carried out using the CHARMM program (Brooks et al., 1983).

Results and Discussion

Table I shows the computed free energy change for the Ile → Val mutation in the fully unfolded (completely solvated) state and for the mutation of residues 16, 61, 78, and 90 in the folded protein. We only performed a single calculation for an Ile → Val unfolded state free energy change since the unfolded state environment is the same for all the mutants in our model, and hence, $\Delta G^{\text{unfolded}}_{\text{native}\rightarrow\text{mutant}}$ for a Val → Ile mutation = $-\Delta G^{\text{unfolded}}_{\text{native}\rightarrow\text{mutant}}$ for an Ile → Val mutation.

The most significant feature of Ile → Val unfolded state free energy change is that it is small in magnitude. This finding is in qualitative agreement with the results of Wolfenden et al. (1981) and the collected data in Ben-Naim and Marcus (1984), which show that the solvation free energy difference between propane and butane in water is 0.2 kcal/mol. While our calculated value does not agree in sign with the experimental findings (i.e., we find butane to be more soluble than propane), it does agree in absolute magnitude, that is, the number is close to zero. If the overall stability change is to have the right magnitude, 1–2.5 kcal/mol, the free energy contribution from mutation in the folded state must dominate.

The free energy change for I61V and I90V mutations in the folded state are 2.3 ± 0.3 and 1.6 ± 0.3 kcal/mol, respectively (Table I). The computed stability changes for mutation, $\Delta\Delta G_{\text{native}\rightarrow\text{mutant}}$, are the differences between these values and the free energy change in the unfolded state and are 1.9 and 1.2 kcal/mol for the I61V and I90V mutations, respectively. The average of these stability changes, 1.5 kcal/mol, corre-

sponds well with the experimental values obtained from mutation studies (Kellis et al., 1988, 1989; Shortle et al., 1990). These results support the view that a loss of favorable interactions in the folded state is the dominant contributor to the observed decrease in stability when isoleucine is mutated to valine.

The simulations enable us to rationalize the stability changes on structural grounds. Figure 2a shows the cavity around Ile-61 before (left) and after (right) mutation. Before mutation, the residues surrounding Ile-61 make favorable van der Waals contacts with the C^{δ} methyl group of the side chain. To quantify this, we have run separate simulations at the isoleucine end point and have computed interaction energies between the isoleucine C^{δ} methyl group and its neighbors (all atoms within 5 Å of C^{δ}). These calculations show that, while the average interaction energy from each favorable van der Waals contact is less than 0.2 kcal/mol, the cavity is efficiently packed such that C^{δ} makes favorable interactions with all of its many neighbors simultaneously. When these relatively small interaction energies are summed over all nearest neighbors, the resulting interaction energy is −1.6 kcal/mol for I61 and −0.8 kcal/mol for I90, each of which constitutes a significant fraction of the folded state free energy change.

The right side of Figure 2a shows the cavity around the side chain after the mutation. The surrounding core residues relax to pack favorably around the smaller side chain, without leaving a cavity. Thus, the free energy change observed on mutation cannot be explained simply on the basis of the energies of leaving a cavity in the protein core (Kellis et al., 1988). Energy calculations at the valine end point show that the valine methyl group $C^{\gamma 1}$ (to which the Ile–C^{δ} methyl was attached in the native protein) forms similar favorable van der Waals interactions with its neighbors (the average interaction energy is −1.1 kcal/mol). Since both the Ile and Val side chains have favorable packing interactions, we ascribe the free energy difference for mutation to the *loss* of a set of net favorable packing interactions with the isoleucine C^{δ} methyl group.

The same structural analysis was performed for the I90V mutant, and very similar behavior was found. In this case, the free energy change for mutation is slightly smaller (Table I), but the character of the packing of the side chain before and after mutation was similar to that of the I61V mutation. The mutation also does not result in a cavity in the interior. Both the isoleucine C^{δ} and the valine C^{γ} methyl groups are favorably packed. For this mutant, we also ascribe the free energy change to the loss of one set of net favorable packing interactions.

While it is observed that most Ile → Val mutations destabilize the folded state, it is not the case that mutations from Val to Ile are stabilizing. From our analysis of the Ile → Val mutations, we found that the core was able to repack around the smaller valine side chain. We have performed two Val → Ile mutations to see whether relaxation in the core can accommodate an extra methylene group. The data in Table I show that the two Val → Ile mutations differ significantly.

The V16I mutation results in a small free energy change of 0.2 kcal/mol in the folded state. The overall stability change, 0.6 kcal/mol, is the same as that measured by Sandberg and Terwilliger (1991) for the V35I mutation in the bacteriophage f1 gene V protein. An inspection of the cavity before and after mutation (Figure 2b) shows that a few simple rearrangements of Val-16 and its neighboring residue, Tyr-4, can result in the formation of a cavity for the additional C^{δ} methylene group of the isoleucine. The energetic cost of these deformations counterbalances the new packing interactions in the Ile side chain, and the overall free energy change in the folded state is close to zero.

In contrast, the V78I mutation results in a large unfavorable change in the free energy of the folded state. The cavities around the side chain before and after mutation are shown in Figure 2c, left and right. The cavity formed around the Ile-78 mutant is more restricted than that formed around the Ile-16 mutant side chain. Apparently, for this mutation the protein cannot relax to form favorable interactions with the larger Ile side chain. The Val-16 side chain points from the α-helix toward the edge of the β-sheet, and has fewer residues separating it from the protein surface (Figure 1). Thus, neighboring groups are free to rearrange around the larger isoleucine side chain at position 16. However, Val-78 lies on the central strand of the β-sheet and points directly toward the body of the helix. In fact, our energetic analysis shows that a C^{γ} methyl group of the Val-78 side chain is favorably packed against the helix backbone. This suggests that the energy change is large for the V78I mutation because the core cannot rearrange without destroying some previously formed favorable interaction. We suspected that the change might be due to disruption in the helical hydrogen bonds, but the excess energy appears to be spread over a large number of interactions and a net destabilization of hydrogen bonds was not obvious from our analysis.

Our assertion that mutation of isoleucine to valine is destabilizing because of a loss of favorable packing interactions supports the argument that the protein core is efficiently packed and that this results in significant stabilization of the folded form of the protein. This argument could explain why Ile → Val mutations are usually destabilizing even though the solvation free energies of Ile and Val differ by only approximately 0.2 kcal/mol [a situation previously discussed by Kellis et al. (1988)]. Thus, we conclude that the stabilization due to burial of the larger isoleucine side chain arises not from an increased aversion of this group for water in the unfolded state but rather from an increase in favorable van der Waals interactions in the core of the folded state. It is possible that the decreased stability of the mutant arises from torsional strain and/or disruption of hydrogen bonds when the core repacks around the smaller side chain. However, our assertion that the loss of favorable packing interactions is the dominant contributor to the observed decrease in stability is consistent with the known relationship between the free energy of transfer of nonpolar solutes transfer from water to nonpolar liquids and the number of carbon atoms being transferred. Experimental solvation data (Ben-Naim & Marcus, 1984; Abraham, 1982) show that the solvation free energy per methylene in water is approximately 0.2 kcal/mol of methylene, while in nonpolar solvents it is approximately −0.7 kcal/mol of methylene (which sum to give −0.9 kcal/mol of methylene, in good agreement with the transfer free energy). Furthermore, these studies indicate that the observed transfer free energy of ∼−0.8 kcal/mol of methylene (Tanford, 1980) arises primarily from favorable enthalpic interactions in the nonpolar liquid (Abraham, 1982).

We are not suggesting that the "hydrophobic effect"[2] in protein folding arises solely from energetic packing interactions, since the free energy for removing any of the nonpolar alkyl side chains from interaction with water is significant: ∼2

[2] The term "hydrophobic effect" has many meanings depending on the context in which it is used (Dill, 1990). In the present context, we use the term to refer to the physical forces that stabilize the folded protein which result from the burial of nonpolar side chains.

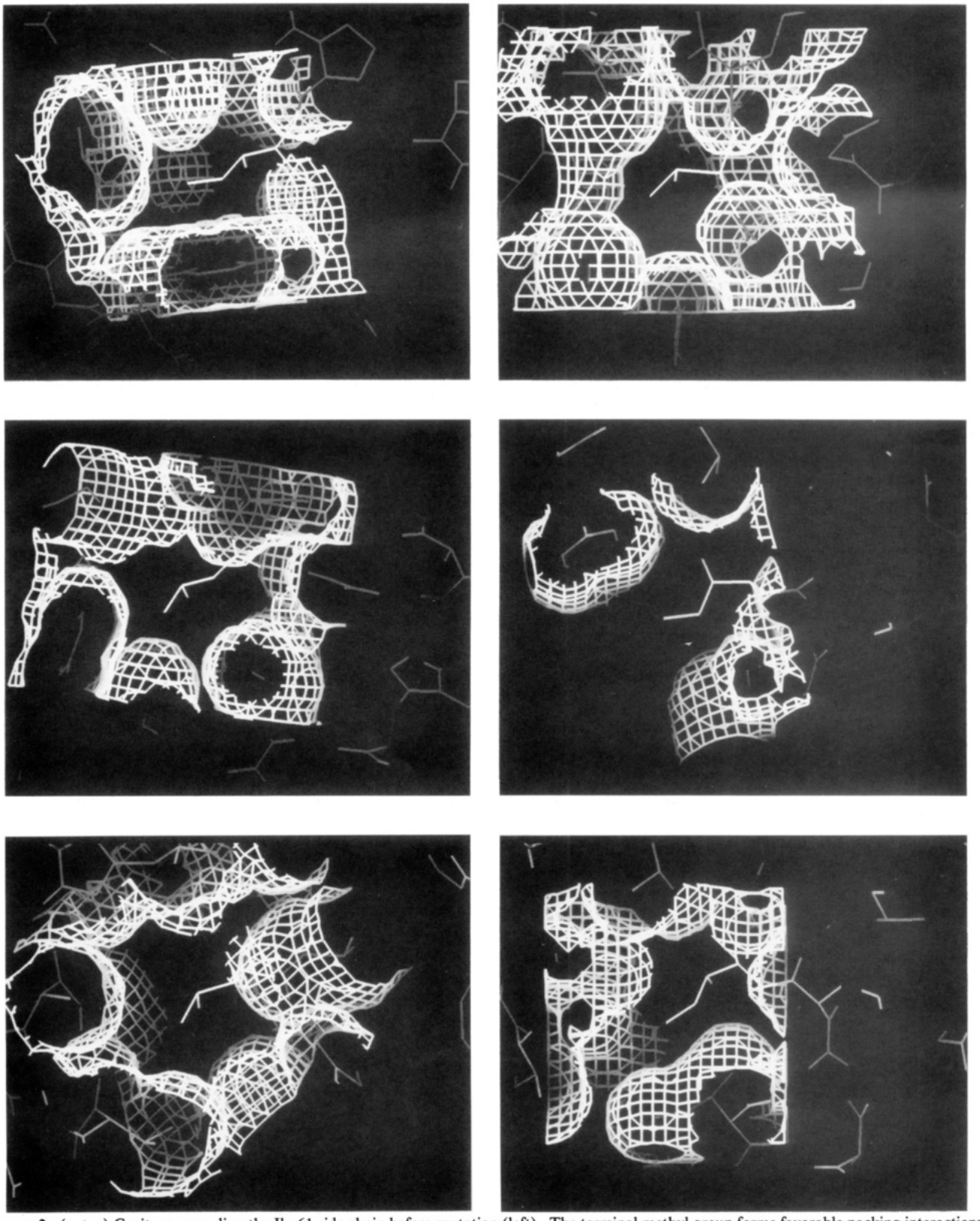

FIGURE 2: (a, top) Cavity surrounding the Ile-61 side chain before mutation (left). The terminal methyl group forms favorable packing interactions with several of its neighbors. The Ile-61 cavity is seen after mutation (right), showing that the core repacks around the smaller side chain; thus, the mutation does not leave a vacancy in the core. A similar repacking of the Val side chain was observed for the I90V mutant. (b, middle) Cavity surrounding the Val-16 side chain before mutations (left). After mutation, the methyl group of isoleucine is accommodated by rotation of the side chain χ angle and motion of the neighboring Try-4 side chain (right). (c, bottom) Cavity surrounding the Val-78 side chain before mutation (left) and after mutation (right). The cavity around the terminal methyl in the Ile-78 mutant appears more restricted than does the cavity for the Ile-16 mutant. There was no simple conformational change which would easily accommodate the extra methylene group (see text).

kcal/mol of side chain (Wolfenden et al., 1981; Ben-Naim & Marcus, 1984). We might term this a "nonspecific hydrophobic effect" because it has roughly the same magnitude for all of the alkyl side chains, and, hence, it is essentially unchanged by mutations. However, once the nonpolar groups are sequestered from interaction with solvent, specific packing

interactions can form, yielding a significant additional stabilization of the folded state which we have found can be on the order of 1.5 kcal/mol of methylene group. The transferability of the value 1.5 kcal/mol of methylene is supported by Ile → Ala and Leu → Ala mutations which typically destabilize the protein by roughly three times the amount as Ile → Val mutation (Kellis et al., 1988, 1989; Shortle et al., 1990; Sandberg & Terwilliger, 1991). This argument about the origin of the stability of the protein core is in accord with the analysis by Murphy et al. (1990), who ascribe the stabilization of the hydrophobic effect to favorable protein–protein interactions which are energetic in origin (e.g., van der Waals). Kellis et al. (1988, 1989) have also interpreted their mutation results in terms of changes in the folded state free energy. Finally, Richards (1977) has also argued that packing interactions were significant on the basis of his investigations of the structure of protein cores.

Our results for the mutation of valine to isoleucine suggest that, while the core can usually relax to fill a cavity (as found in the Ile → Val mutations), it cannot readily expand to accommodate an extra methylene group. While this certainly depends on the exact location of the mutation, it suggests to us that not only is the core efficiently packed but it is also almost completely full in RNase T_1. Thus, the addition of extra methylene groups will usually occur at the expense of existing favorable interactions (e.g., packing and/or optimal hydrogen bonding).

Registry No. Ile, 73-32-5; Val, 72-18-4; ribonuclease T_1, 9026-12-4.

REFERENCES

Abraham, M. H. (1982) *J. Am. Chem. Soc. 104*, 2085–2094.

Arni, R., Heinemann, U., Tokuoka, R., & Saenger, W. (1988) *J. Biol. Chem. 263*, 15358–15368.

Behe, M. J., Lattman, E. E., & Rose, G. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4195–4199.

Ben-Naim, A., & Marcus, Y. (1984) *J. Chem. Phys. 81*, 2016–2027.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983), *J. Comput. Chem. 4*, 187–217.

Dang, L. X., Merz, K. M., Jr., & Kollman, P. A. (1989) *J. Am. Chem. Soc. 111*, 8505–8508.

Dill, K. A. (1990) *Biochemistry 29*, 7133–7155.

Fleischman, S. H., & Brooks, C. L., III (1987) *J. Chem. Phys. 87*, 3029–3037.

Gellman, S. H. (1991) *Biochemistry 30*, 6633–6636.

Hermans, J., Yun, R.-H., & Anderson, A. G. (1990) in *Crystallographic and Modelling Methods in Molecular Design* (Bugg, C. E., & Ealick, S. E., Eds.) Springer-Verlag, New York.

Hill, C., Dodson, G., Heinemann, U., Saenger, W., Mitsui, Y., Nakamura, K., Borisov, S., Tischenko, G., Polyakov, K., & Pavlovsky, S. (1983) *Trends Biochem. Sci. 8*, 364–369.

Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.

Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W., & Klein, M. L. (1983) *J. Chem. Phys. 79*, 926–935.

Kauzmann, W. (1959) *Adv. Protein Chem. 14*, 1–63.

Kellis, J. T., Jr., Nyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature 333*, 784–786.

Kellis, J. T., Jr., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914–4922.

Lim, W. A., & Sauer, R. T. (1989) *Nature 339*, 31–36.

Murphy, K. P., Privalov, P. L., & Gill, S. J. (1990) *Science 247*, 559–561.

Pace, C. N. (1990) *Trends Biochem. Sci. 15*, 14–17.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327–341.

Sandberg, W. S., & Terwilliger, T. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1706–1710.

Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033–8041.

Tanford, C. (1980) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, John Wiley and Sons, New York.

Tidor, B., & Karplus, M. (1991) *Biochemistry 30*, 3217–3228.

Verlet, L. (1967) *Phys. Rev. 159*, 98–103.

Wolfenden, R., Andersson, L., Cullis, P. M., & Southgate, C. C. B. (1981) *Biochemistry 22*, 849–855.

Wong, C. F., & McCammon, J. A. (1986) *J. Am. Chem. Soc. 108*, 3830–3832.