Protein structure, stability and solubility in water and other solvents

C. Nick Pace[*], Saul Treviño, Erode Prabhakaran and J. Martin Scholtz

Department of Medical Biochemistry and Genetics, Department of Biochemistry and Biophysics and Center for Advanced Biomolecular Research, Texas A&M University, College Station, TX 77843, USA

Proteins carry out the most difficult tasks in living cells. They do so by interacting specifically with other molecules. This requires that they fold to a unique, globular conformation that is only marginally more stable than the large ensemble of unfolded states. The folded state is stabilized mainly by the burial and tight packing of over 80% of the peptide groups and non-polar side chains. If life as we know it is to exist in a solvent other than water, the folded state must be stable and soluble in the new solvent. Our analysis suggests that proteins will be unstable in most polar solvents such as ethanol, extremely stable in non-polar solvents such as cyclohexane, and even more stable in a vacuum. Our solubility studies suggest that protein solubility will be markedly lower in polar solvents such as ethanol and that proteins will be essentially insoluble in non-polar solvents such as cyclohexane. For these and other reasons it seems unlikely that the life we know could exist in any solvent system other than water.

Keywords: protein structure; protein stability; hydrophobic bonds; hydrogen bonds; protein folding; protein solubility

1. INTRODUCTION

Proteins do the most demanding jobs in living cells. To do so, they must interact specifically with other molecules. This requires that the polypeptide chain fold to a unique, globular conformation that we will call the native state of the protein and denote by N. This is the structure that can be determined in crystals by X-ray crystallography or in solution by NMR. The native state is only marginally more stable—generally 2–10 kcal mol^{-1} under physiological conditions—than the large ensemble of unfolded states that we will call the denatured state of the protein and denote by D. We can represent the equilibrium between these two macrostates as

$$N \leftrightarrow D \tag{1.1}$$

and define the conformational stability of a protein as

$$\Delta G = G_D - G_N = -RT \ln K = -RT \ln[D]/[N], \tag{1.2}$$

where [D] and [N] represent the concentrations of D and N, G_D and G_N represent the free energies of D and N, and K and ΔG are the equilibrium constant and standard free energy change, respectively, for equation (1.1). Several methods are available for measuring the conformational stability of a protein or the difference in stability between the wild-type protein and a variant differing by a single amino acid (Pace & Scholtz 1997).

RNase Sa is a small globular protein in the microbial ribonuclease family. Figure 1 shows a schematic of the folding of RNase Sa. Much of our laboratory information

about proteins has come from studies of RNase Sa so we will use it as an example in this article (Takano *et al.* 2003). The crystal structure of the protein has been determined at 1.0 Å resolution by Sevcik *et al.* (2002) and a solution structure has been determined using NMR by Laurents *et al.* (2001). In addition, the p*K* values of the ionizable groups were recently measured (Huyghues-Despointes *et al.* 2003; Laurents *et al.* 2003). Two other members of the microbial RNase family have been studied in detail: RNase T1 (Pace *et al.* 1991; Steyaert 1997; Giletto & Pace 1999) and barnase (Hebert *et al.* 1998; Khan *et al.* 2003).

It is clear that if life as we know it is to exist in a solvent other than water two important conditions must be met: first, the native state of the protein must be favoured, that is, the conformational stability must be 2–10 kcal mol^{-1} in the new solvent; and, second, surely N, and perhaps D, must be soluble in the new solvent. First, we will summarize what is known about the structure, stability and solubility of the N and D states in water. Next, we will consider what happens to the structure, stability and solubility of a protein in a vacuum, in a non-polar solvent such as cyclohexane and in a polar solvent such as ethanol.

2. THE NATIVE STATE

Figure 2 is a schematic of protein folding from Tanford (1962). It illustrates two key points: first, when a protein unfolds, many peptide groups and side chains that are buried in N become exposed to solvent in D; and, second, there may be pockets of structure in D. Table 1 shows that 83% of the most non-polar side chains and 82% of the peptide groups are buried in N (Lesser & Rose 1990). Consequently, these are the most important groups that become exposed to solvent when a protein unfolds. We

[*] Author for correspondence (nickpace@tamu.edu).

Phil. Trans. R. Soc. Lond. B (2004) **359**, 1225–1235
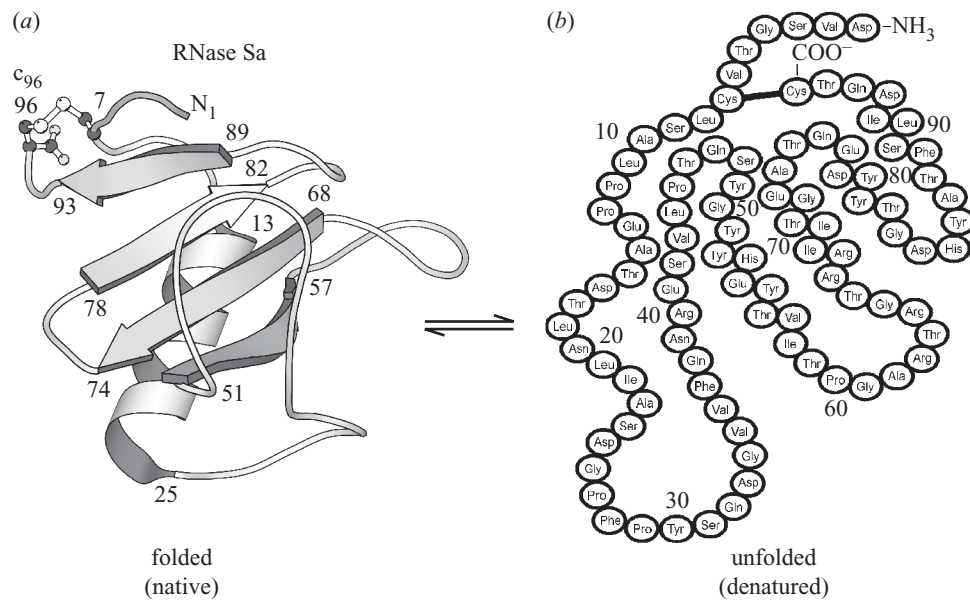DOI 10.1098/rstb.2004.1500

1225

Figure 1. Schematic of the folding of RNase Sa. A ribbon diagram based on a 1.2 Å resolution crystal structure determined by Sevcik *et al.* (1996) is used to illustrate (*a*) the native state (N), and (*b*) the amino acid sequence is used to represent the denatured state (D). The difference in free energy between these two states defines the conformational stability of the protein (equations (1.1) and (1.2)). At pH 7, the conformational stability of RNase Sa is 6.1 kcal mol^{-1} and the melting temperature $T_m = 48.4\,°C$. When the single disulphide bond is broken, the protein still folds, but the conformational stability is lowered by *ca.* 5 kcal mol^{-1} and $T_m = 28.4\,°C$ (Pace *et al.* 1998).
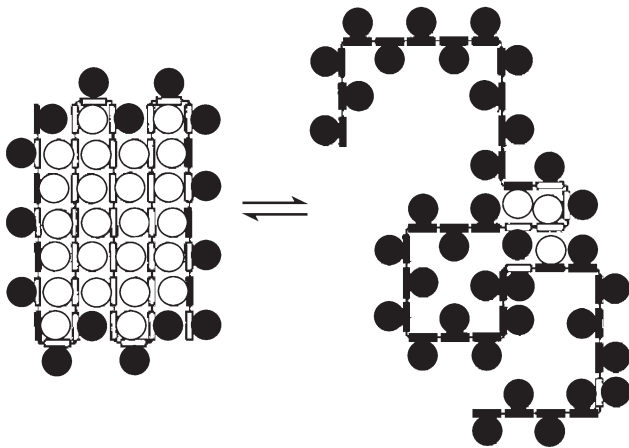

Figure 2. Schematic of protein folding (Tanford 1962). This illustrates that many of the groups buried in the native state are accessible to solvent after the protein unfolds. Table 1 shows that the groups exposed to solvent after unfolding are mainly peptide groups and non-polar side chains. Table 2 lists the free energy of transfer, ΔG_{tr}, for peptide groups and a non-polar leucine side chain from water to various solvents. By considering these data it is easy to see why urea and GdnHCl are protein denaturants and why trimethylamine oxide (TMAO) and sarcosine are protein stabilizers (osmolytes).

Table 1. Folded globular proteins.

	buried[a]
non-polar side chains (Ala, Val, Ile, Leu, Met, Phe, Trp, Cys)	83%
peptide groups (−CO–NH−)	82%
ca. 1.1 intramolecular hydrogen bonds formed per residue[b]	

[a] Lesser & Rose (1990).
[b] Stickle *et al.* (1992).

Table 2. ΔG_{tr} (cal mol^{-1}) from water to solvent.

solvent	peptide group	Leu side chain
urea (2 M)[a]	−70	−110
GdnHCl (2 M)[a]	−135	−210
sarcosine (2 M)[b]	+90	+80
trimethylamine oxide (2 M)[b]	+180	+20
ethanol[c]	+1400	−1800
cyclohexane[d]	+7600	−4900
vacuum[e]	+9800	−2300

[a] Pace (1975); Liu & Bolen (1995).
[b] Bolen & Baskakov (2001).
[c] Nozaki & Tanford (1971).
[d] Radzicka & Wolfenden (1988).
[e] Privalov & Makhatadze (1993).

will be interested in how much each of these contributes to the stability of the protein. In addition, 1.1 intramolecular hydrogen bonds are formed per residue when a protein folds (Stickle *et al.* 1992). The non-polar side chains and peptide groups that are buried in the folded protein are tightly packed. This is illustrated in table 3, which compares the packing in three solvents with the packing of close-packed spheres and groups in the interior of a protein. On this basis, Klapper (1971) concluded 'the protein interior contains little space and is closer to a solid than a liquid'. As discussed below in §§ 5c,d, we think this feature of protein structure is of crucial importance to protein stability.

Table 3. *Protein interiors are tightly packed*[a].

compound	space occupied by atoms (%)
water	0.36
cyclohexane	0.44
ethanol	0.47
close-packed spheres	0.71
protein interior	0.75

[a] Klapper (1971).

Table 4. *Percentage burial of surface for proteins with 100 and 700 amino acids*[a].

	percentage burial	
	100-residue protein	700-residue protein
aliphatic	57	77
aromatic	73	83
polar uncharged	57	78
polar charged	37	61

[a] Kajander *et al.* (2000).

As globular proteins become larger, they must bury a greater percentage of their residues. This is illustrated by the results in table 4, which show, surprisingly, that when proteins become larger there is a larger increase in the burial of charged groups than for any other type of group. The authors suggest that this may be an approach that evolution uses to keep proteins from becoming too stable (Kajander *et al.* 2000). However, for both smaller and larger proteins, the groups most likely to be accessible to solvent on the native state of a protein are the charged groups. The proportion of the surface that is composed of charged side chains is the most important determinant of the solubility of a protein (Malissard & Berger 2001).

3. THE DENATURED STATE

Since the first high-resolution crystal structures became available in the early 1960s, our detailed view of the structure of the native states of proteins has increased greatly, but the broad view has not changed. By contrast, our understanding of the denatured state has changed over the years and is still in a state of flux (Eisenberg *et al.* 2002; Shortle 2002). Tanford's group carried out the first careful studies of the denatured states of proteins and concluded that proteins approach a randomly coiled conformation in 6 M GdnHCl with their disulphide bonds broken (Tanford 1968). Similar studies reached the same conclusion for proteins in 8 M urea. It was clear to Tanford that pockets of structure might exist, and this was emphasized at the time by polymer chemists (Miller & Goebel 1968). More recent studies of the denatured state using small-angle X-ray scattering to determine the radius of gyration, R_g, or pulsed-field-gradient NMR to estimate the hydrodynamic radius, have reached similar conclusions (Millet *et al.* 2002).

Table 5. *Solubility of folded and unfolded RNase Sa in different solvents*[a].

protein and solvent	solubility (mg ml^{-1})
folded in H$_2$O at 25 °C	2.05 ± 0.11
unfolded in H$_2$O at 65 °C[b]	0.01 ± 0.01
unfolded and reduced in H$_2$O at 25 °C[c]	< 0.01 ± 0.01
folded in H$_2$O at 25 °C	2.05 ± 0.11
folded in 20% ethanol at 25 °C	0.09 ± 0.01
folded in 2 M urea at 25 °C	12.06 ± 1.20

[a] Solubility measured between pH 3.5 and 4.0, near the pI.
[b] Unfolded RNase Sa with the disulphide bond intact at 65 °C.
[c] Unfolded RNase Sa with the disulphide bond broken at 25 °C.

In a recent study, Goldenberg (2003) generated ensembles of the denatured state that were restricted only by the excluded volume of the protein. By comparing his results to experimental data, he concluded that '…the overall properties of unfolded proteins can be usefully described by a random coil model and that an unfolded polypeptide can undergo significant collapse while losing only a relatively small fraction of its conformational entropy' (p. 1615). Figure 3 is taken from the Goldenberg (2003) paper and it makes another important point. The effect of excluded volume is to give compact denatured states with R_g values much closer to the native protein than to a protein in an extended conformation. Nevertheless, the solvent accessibility of the denatured state ensemble is much closer to that of a protein in an extended conformation than to that of the native protein. The fact that the conformational entropy and solvent accessibility are both high for compact denatured states is important in understanding protein stability.

Proteins are unfolded to the greatest extent in urea and GdnHCl solutions, but it is clear that hydrophobic clusters and other native-like structures exist even under strongly denaturing conditions (Denisov *et al.* 1999). NMR is beginning to give a clearer picture of the denatured state ensemble that exists under physiological conditions, the denatured state of most interest to us. Choy *et al.* (2002) have shown that the molecules in the denatured state ensemble are relatively compact, *ca.* 30–40% larger than the native state. Mayor *et al.* (2003) studied the denatured state of the engrailed homeodomain and concluded: 'The denatured state had extensive native secondary structure and was significantly compact and globular. But, the side-chains and backbone were highly mobile' (p. 977). It has become clear in recent years that some proteins are unfolded even under physiological conditions and these are referred to as 'natively unfolded' proteins (Uversky 2002).

4. SOLUBILITY OF THE NATIVE AND DENATURED STATES

RNase Sa is an acidic protein with an isoelectric pH (pI) of 3.5 that contains no Lys residues. By replacing Asp and Glu residues on the surface of RNase Sa with Lys residues, we created a variant with a pI of 6.4 and a variant
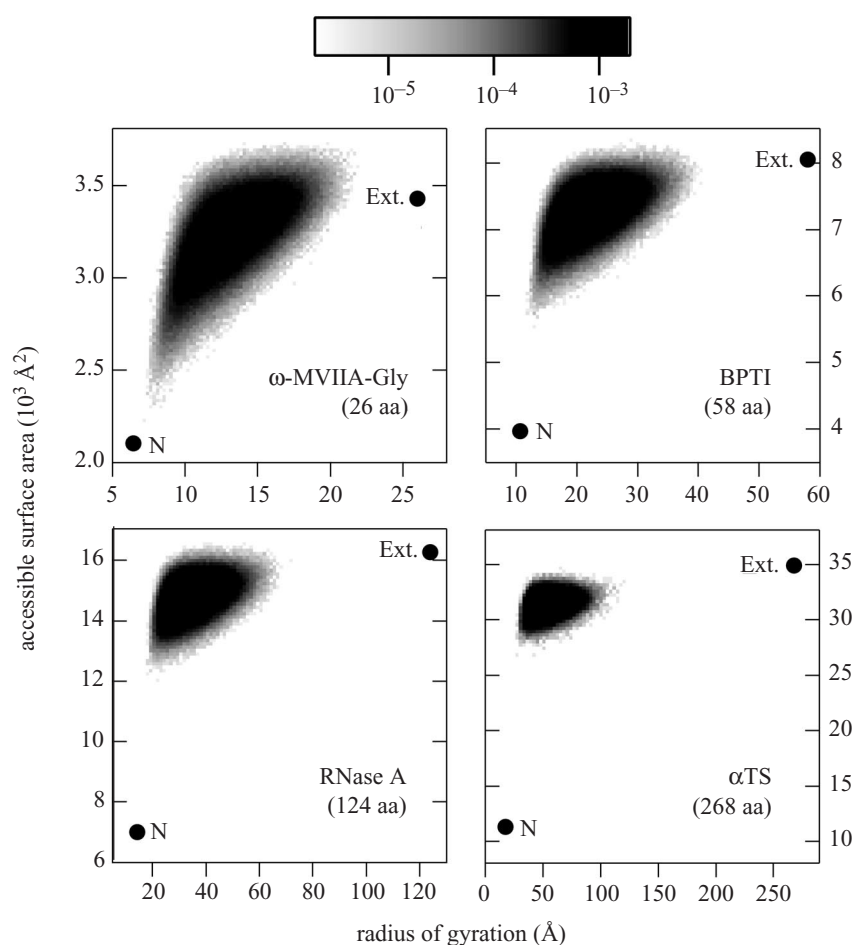
Figure 3. Distributions of radius of gyration and accessible surface area for ensembles of conformations for four proteins simulated with excluded volume (Goldenberg 2003). The scale bar shows the fraction of the total population of chains for which the radius of gyration and accessible surface area fall within a given area. Filled circles denote the accessible surface area and radius of gyration for the native protein, denoted N, and the protein in an extended conformation, denoted Ext. ω-MVIIA-Gly, ω-conotoxin MVIIA-Gly; BPTI, bovine pancreatic trypsin inhibitor; RNase A, bovine pancreatic ribonuclease A; αTS, α-subunit of *Escherichia coli* tryptophan synthetase.

with a pI of 10.2 (Shaw *et al.* 2001). We showed that the minimum solubility of the protein was always near the pI, as has been observed previously (Cohn & Edsall 1943). We wondered how much lower the solubility of the denatured state would be under the same conditions. To study this, we reduced the disulphide bond in RNase Sa, which lowers the melting temperature so that the denatured state can be studied at low pH at 25 °C. The results in figure 4 show that the solubility of the denatured state was too low to measure near the pI. Note that the solubility of native RNase Sa increases sharply as the net charge on the protein becomes either positive or negative. By contrast, a larger charge seems to be needed on the denatured state to solubilize the protein. We also measured the solubility of the thermally denatured protein with the disulphide bond intact at 65 °C. The results are shown in table 5. Again, the denatured protein is much less soluble than native RNase Sa, but the solubility is not as low as it was for reduced, denatured RNase Sa at 25 °C.

To gain an understanding of how much less soluble a protein would be in non-polar solvents, we measured the solubility of RNase Sa in 20% ethanol where the protein is folded. The solubility is decreased more than 20-fold by the presence of 20% ethanol (table 5). On this basis,

it seems likely that most proteins will be almost completely insoluble in most organic solvents, at least near their pI. Given that all of the constituent groups in a protein are more soluble in urea solutions than they are in water, we would expect urea to increase the solubility of proteins (Qu *et al.* 1998). The presence of 2 M urea does indeed increase the solubility of RNase Sa almost sixfold compared with water (table 5).

5. FORCES STABILIZING PROTEINS

(a) *Conformational entropy*

The major force destabilizing proteins is conformational entropy. Rotation around the many bonds in a protein is much freer in the denatured state than in the native state and provides a strong entropic driving force for protein unfolding. As discussed previously (Pace *et al.* 1998), we have used approaches developed by Spolar & Record (1994) and by D'Aquino *et al.* (1996) to estimate that conformational entropy favours the denatured state of RNase Sa by *ca.* 165 kcal mol^{-1} at 25 °C. We will now consider the forces that are most important in stabilizing the native state.
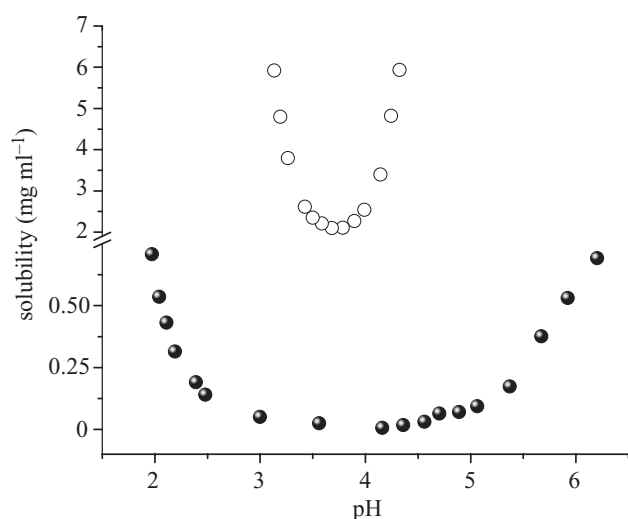
Figure 4. Solubility of native (open circles) and denatured (filled circles) RNase Sa as a function of pH at 25 °C. Denatured RNase Sa was prepared by reducing the disulphide bond using tricarboxyethylphosphine for 60 min at 60 °C.

(b) *Charge–charge interactions and salt bridges*

The charges on the surface of proteins are generally arranged so that there are more attractive than repulsive interactions near neutral pH (Karshikoff & Ladenstein 1998). Consequently, these electrostatic interactions will generally contribute favourably to a protein's stability. Nevertheless, several recent studies have shown that it is possible to stabilize proteins by making charge reversal mutations on the surface that improve the electrostatic interactions even further (e.g. Grimsley *et al.* 1999). Such stability increases are always less than predicted by using Coulomb's law and a dielectric constant of 80 to sum up the electrostatic interactions on the native state. This led us to conclude that the charge–charge interactions that stabilize the native states of proteins also contribute favourably to the denatured states so that the net contribution to protein stability is small (Pace *et al.* 2000). Thus, it is unlikely that charge–charge interactions will make contributions to the protein stability of greater than 10 kcal mol^{-1} at 25 °C. At higher temperatures, the contribution might be considerably greater, and it appears that the proteins from thermophilic organisms often use this strategy to increase their stability (Elcock 1998; de Bakker *et al.* 1999; Xiao & Honig 1999; Pace 2000).

When oppositely charged groups on the surface of a protein are within 5 Å, they are generally referred to as ion pairs or salt bridges. (Sometimes only ion pairs that are close enough to form hydrogen bonds are referred to as salt bridges.) Ion pairs on the surface of a protein generally contribute less than 1 kcal mol^{-1} to the stability (Marti & Bosshard 2003). However, a buried salt bridge can contribute more than 4 kcal mol^{-1} to the stability (Anderson *et al.* 1990), but the number of buried salt bridges in any given protein is small so they do not make a large contribution to protein stability.

At pH 7, RNase Sa has seven positive charges and 13 negative charges so that the net charge is −6. If we use Coulomb's law with a dielectric constant of 80 and sum up the charge–charge interactions for native RNase Sa at pH 7, it leads to an estimate that electrostatic interactions destabilize the protein by *ca.* 5 kcal mol^{-1}. In the denatured state ensemble, the charges will, on average, probably be more favourably arranged so the repulsion will be less. Thus, electrostatic interactions probably do not make a large contribution to the stability of RNase Sa. In water, electrostatic interactions will generally make a smaller contribution to the stability than the hydrophobic effect and hydrogen bonding, but they can become a dominant force in a vacuum or a non-polar solvent.

(c) *Hydrophobic interactions*

In an influential review, Dill (1990) concluded:

> More than 30 years after Kauzmann's insightful hypothesis, there is now strong accumulated evidence that hydrophobicity is the dominant force of protein folding.... There is evidence that hydrogen bonding or van der Waals interactions among polar amino acids may be important, but their magnitude remains poorly understood.
>
> (Dill 1990, p. 7151)

Most biochemists still believe that hydrophobicity is the dominant force in protein folding, but we think it has become clear in the years since Dill's review that hydrogen bonding and van der Waals interaction of polar groups make a contribution to protein stability comparable to that from hydrophobicity. The evidence will be presented in § 5d. The review of Kauzmann (1959) presented convincing evidence that hydrophobic interactions make a major contribution to protein stability. Tanford (1962) went even further and concluded: '...the stability of the native conformation in water can be explained...entirely on the basis of the hydrophobic interactions of the non-polar parts of the molecule' (p. 4245).

On the basis of experimental studies of hydrophobic interactions in proteins, the burial of a −CH$_2$− contributes *ca.* 1.2 kcal mol^{-1} or 49 cal mol^{-1} per Å^3 to the stability of a protein (Pace 2001). For comparison, the ΔG_{tr} for transfer of a −CH$_2$− from water to cyclohexane is *ca.* 1.0 kcal mol^{-1} or 39 cal mol^{-1} per Å^3 (Pace 1995). These results suggest that *ca.* 80% of the hydrophobic effect is due to hydrophobicity and the other 20% is due to the tight packing of hydrophobic groups in the protein interior. Using a different approach, Chen & Stites (2001) concluded that 'Hydrophobicity has long been thought to be the major driving force for protein stability. Recently, close packing of the hydrophobic core to optimize van der Waals contacts and minimize cavities has been proposed to be of roughly equal importance energetically' (p. 442).

For RNase Sa, the contribution of the hydrophobic effect to the stability would be 88 kcal mol^{-1} if n-octanol is used as a model for the interior of a protein, 150 kcal mol^{-1} if cyclohexane is used as a model for the interior of the protein, and 160 kcal mol^{-1} if we use the best estimates based on experimental studies of proteins (Pace 1995). Each of these estimates is likely to be too high since we estimate the extent of burial of the non-polar groups based on a model of the unfolded state with an accessible surface area that is too high. However, these estimates show clearly that the hydrophobic effect does make a major contribution to protein stability.

Table 6. Volume, hydrophobicity and side chain conformational entropy changes for Tyr → Phe and Thr → Val mutations.

	Tyr → Phe	Thr → Val
volume[a]	$-5\,\text{Å}^3$	$18\,\text{Å}^3$
hydrophobicity[b]	$1.13\,\text{kcal mol}^{-1}$	$1.31\,\text{kca mol}^{-1}$
side-chain entropy[c]	$-0.51\,\text{kcal mol}^{-1}$	$-0.65\,\text{kcal mol}^{-1}$

[a] Based on residue volumes in proteins (Tsai *et al.* 1999).
[b] Based on the ΔG_{tr} values for n-octanol (Fauchere & Pliska 1983).
[c] Based on the mean $T\Delta S$ values at 300 K (Doig & Sternberg 1995).

(d) *Hydrogen bonding and polar group burial*

In 1936, Mirsky & Pauling (1936) concluded: 'The importance of the hydrogen bond in protein structure can hardly be overemphasized' (p. 15 280). This view was strengthened in the early 1950s when the α-helix and β-sheet were proposed as major structural elements in proteins by Pauling *et al.* (1951). Kauzmann's well-reasoned review (Kauzmann 1959) changed most biochemists' minds and the hydrophobic effect was viewed as the dominant force in protein folding. This is the view that prevails today.

The first good evidence that hydrogen bonds might contribute favourably to protein stability came from studies of the interaction of tRNA synthetase with its substrates (Fersht 1987). They showed that hydrogen bonds generally contribute more than 1 kcal mol^{-1} per hydrogen bond to the energetics of substrate binding. Since that time, many other experimental studies have reached similar conclusions (Myers & Pace 1996).

To gain a better understanding of the contribution of hydrogen bonding to protein stability, we have made mutants that add or remove side chains capable of forming hydrogen bonds and measure their stability to compare with the wild-type protein. The three mutations we will discuss here are Tyr → Phe, Thr → Val and Val → Thr (Pace *et al.* 2001; Takano *et al.* 2003). Table 6 gives some information characterizing these mutations. The differences in hydrophobicity and side-chain conformational entropy would both be expected to increase the stability of the Tyr → Phe and Thr → Val mutants. Consequently, if the contribution of hydrogen bonds to the stability were energetically neutral, we would expect the Tyr → Phe mutants to be *ca.* 1.6 kcal mol^{-1} and the Thr → Val mutants *ca.* 2.0 kcal mol^{-1} more stable than the wild-type protein.

In the Tyr → Phe mutants, an −OH group is removed from the protein. This could leave a small cavity in the protein, but should have little effect on the conformation of the protein (Matthews 1995). We have studied 20 Tyr → Phe mutations in three different microbial RNases (Shirley *et al.* 1992; Pace *et al.* 2001). These results, together with results from other groups, are summarized in table 7. On average, the Tyr → Phe mutants were 1.4 kcal mol^{-1} less stable than wild-type when the Tyr−OH group was hydrogen bonded and 0.2 kcal mol^{-1} less stable when it was not. This shows clearly that hydrogen bonds contribute favourably to protein stability. Thus,

the hydrogen bonding and van der Waals interactions of the −OH group in a folded protein are more favourable than the interactions with water in the unfolded protein. Even, when the −OH group is not hydrogen bonded, it makes a favourable contribution to the stability. This shows that the theoretical calculations that '...suggest that polar groups prefer to be fully solvated in water rather than hydrogen bonded in the interior of a protein...' (Honig 1999, p. 286) are wrong.

In the Thr → Val mutants, an −OH group is replaced by a −CH$_3$ group that is slightly larger so that the mutant might be destabilized by steric strain. Again, the mutants are less stable when the −OH group is replaced if it is hydrogen bonded, showing that the hydrogen-bonded −OH group is making a larger contribution to the stability than can be gained by burying a −CH$_3$ group. Even when the Thr −OH group is not hydrogen bonded, the buried −OH group makes a contribution to the stability equivalent to that of burying a −CH$_3$ group at the same site.

We also considered Val → Thr mutations. When the Val −CH$_3$ group is replaced by an −OH group, the average decrease in stability is 1.8 kcal mol^{-1} and it is unfavourable in 39 out of 40 mutants (Takano *et al.* 2003). We expect a decrease in stability of *ca.* 2 kcal mol^{-1} for this substitution (table 6). This shows that replacing a buried −CH$_3$ group with an −OH group at a site designed for a −CH$_3$ group is almost always unfavourable.

These results show that the contribution of polar group burial to protein stability is strongly context dependent. Burying non-polar groups generally stabilizes the native state of proteins. Burying polar groups stabilizes the native state only when the site was designed to accommodate the polar group. Favourable van der Waals interactions and longer-range electrostatic interactions are just as important as hydrogen bonds in stabilizing buried polar groups.

We reached a similar conclusion using results from different mutants and a different approach (Pace 2001). On the basis of studies of Asn → Ala mutants, the gain in stability from burying amide groups that are hydrogen bonded to peptide groups is *ca.* 80 cal mol^{-1} per Å^3. Similar studies of Leu → Ala mutants showed that the gain in stability from burying −CH$_2$− groups is 50 cal mol^{-1} per Å^3. Harpaz *et al.* (1994) analysed the completely buried residues in a sample of 108 proteins. The total volume of completely buried residues was 298 100 Å^3. Of this, 118 200 Å^3 was occupied by non-polar side chains and 92 000 Å^3 was occupied by peptide groups. On this basis, the burial of non-polar side chains contributes 5800 kcal mol^{-1} to the stability and the burial of peptide groups contributes 7200 kcal mol^{-1} to the stability. This amounts to a contribution of 59 kcal mol^{-1} per protein from non-polar group burial and 74 kcal mol^{-1} per protein from peptide group burial. This is further support that polar group burial makes a favourable contribution to protein stability and it might even be greater than the contribution of non-polar group burial.

In summary, all mutational studies of proteins support the idea that hydrogen bonding and polar group burial make a favourable contribution to protein stability. This is supported by recent analyses (Guerois *et al.* 2002; Lomize *et al.* 2002) that are based on results from experimental studies of hundreds of mutants for which the stability was

Table 7. $\Delta(\Delta G)$ values for 52 Tyr → Phe mutants and 40 Thr → Val mutants.

	hydrogen bonded		not hydrogen bonded	
mutation	number	$\Delta(\Delta G)$ (kcal mol^{-1})	number	$\Delta(\Delta G)$ (kcal mol^{-1})
Tyr → Phe[a]	35	-1.4 ± 0.9	17	-0.2 ± 0.4
Thr → Val[b]	25	-0.9 ± 1.0	15	0.0 ± 0.5

[a] Fifty-two Tyr → Phe mutants (Pace *et al.* 2001).
[b] Forty Thr → Val mutants (Takano *et al.* 2003).

measured. Beginning with Brandts (1964) and continuing to the present (Cooper 2000; Loladze *et al.* 2002), it has been clear that the results from studies of the thermodynamics of protein denaturation support the idea that hydrogen bonds and the van der Waals interactions of polar groups must make a large contribution to the enthalpy change for protein folding. In addition, results from a number of other experimental studies using a variety of different approaches also support this idea (Myers & Pace 1996). Nevertheless, even now most biochemists still regard hydrophobic interactions as the dominant force in protein folding. One reason for this is that most theoretical studies of the contribution of hydrogen bonding to protein stability have come to a different conclusion from the experimental studies. We think the reason for this is that the theoretical studies are not properly estimating the contribution of dispersion forces to the contribution of polar group burial to protein stability.

6. PROTEINS IN A VACUUM, IN CYCLOHEXANE, AND IN ETHANOL

We will now use the ΔG_{tr} values given in table 2 to predict what would happen to a folded globular protein when it is transferred from water to ethanol, cyclohexane or a vacuum. In ethanol, the contribution of peptide group burial to protein stability would be enhanced and the contribution of non-polar group burial would be diminished. Since about equal amounts of the two types of groups are buried (table 1), the ΔG_{tr} values suggest that the protein would unfold in ethanol. This is what is generally observed, but the concentration of ethanol required to cause the protein to unfold varies among proteins (Tanaka *et al.* 2001). After a protein is unfolded in ethanol, it is often observed to refold into rod-like structures with a high content of α-helices (Hirota-Nakaoka & Goto 1999). This is not surprising. By folding into α-helices, the protein is able to bury its peptide groups in the interior of the helix out of contact with the ethanol and leave its non-polar side chains exposed to ethanol. Thus, α-helices are expected to be stable structures in ethanol. This is at least part of the reason that compounds such as trifluoroethanol are frequently used to increase α-helix formation in peptides and proteins (Luo & Baldwin 1997).

In cyclohexane, the ΔG_{tr} values go in the same direction as in ethanol, but are much larger (table 2). The peptide groups will have favourable van der Waals interactions with the cyclohexane, but no hydrogen bonding will be possible so they would much prefer to be in water. The non-polar side chains have a strong preference for cyclohexane over water. Thus, these data suggest that folded globular proteins will be very stable in cyclohexane. The model compound data do not take into account the much tighter packing of groups in folded proteins compared with the packing when the protein is unfolded. This will tend to make the folded conformation even more stable in cyclohexane.

Over the past 15 years there has been great interest in using enzymes to catalyse chemical reactions in non-polar solvents (Klibanov 2001). Enzymes are not soluble in these solvents but form suspensions with the enzyme retaining some water molecules on the surface. For example, when subtilisin is crystallized from dioxane, it has 65 bound water molecules but has a conformation similar to the enzyme crystallized from water or acetonitrile (Schmitke *et al.* 1997). The stability of RNase A has been studied in the anhydrous alkane nonane. Under conditions where $T_m = 61$ °C in water, $T_m = 124$ °C in nonane. This is an enormous increase in stability. Klibanov (2001) concluded: 'These and similar data indicate that enzymes are predictably extremely thermostable in anhydrous organic solvents owing to their conformational rigidity in the dehydrated state...' (p. 243). We think that increased rigidity in the native state would destabilize the enzyme because conformational entropy would now favour the denatured state to an even greater extent. Instead, we think the greater stability is due to the much greater gain in stability from the burial of peptide groups in the native protein when the protein is in a non-polar solvent rather than in water (table 2). (See Klibanov (2001) for a consideration of the kinetic barrier to unfolding in organic solvents.)

In a vacuum, the ΔG_{tr} values for peptide groups are even more unfavourable than for cyclohexane and the ΔG_{tr} values for non-polar groups are less favourable than for cyclohexane (table 2). The peptide groups markedly prefer water because no hydrogen bonds or van der Waals interactions will be possible in a vacuum. The non-polar groups prefer a vacuum over water because of the unfavourable hydrophobic effect observed when non-polar groups are added to water. Thus, this analysis predicts that proteins will be much more stable in a vacuum than in cyclohexane, and in both cases much more stable than they are in water.

It was surprising to learn that proteins could remain folded in a vacuum (Wolynes 1995). However, from the previous discussion, it should have been expected. In a vacuum, proteins remain folded until the net charge is large enough that unfavourable charge–charge interactions overcome the stabilizing interactions and cause the protein

to unfold. Techniques have been developed to probe the conformation of proteins in a vacuum. On the basis of studies of cytochrome *c* using these approaches, Jarrold (1999) concluded: 'The +3 to +5 charge states have cross-sections that are slightly smaller than expected for the native conformation suggesting that the protein packs more tightly in the absence of the solvent' (p. 362). It is hard to imagine that proteins can pack much more tightly than they do in water (table 3). It is also doubtful that they could adopt a new folded conformation that is more tightly packed than the native state present in water. However, there could be small conformational changes that would relieve some of the charge repulsion in the native state. After the protein unfolds, the cross-section appears to increase approximately linearly as the net charge increases (Jarrold 1999). We have to guess what conformation the protein will adopt. Several forces would be in operation. The protein would extend as much as possible to minimize unfavourable charge–charge interactions in the denatured state. However, if it could 'self solvate' to reduce the net charge the protein might do so. Molecular dynamics studies suggest that the protein does this by forming hydrogen bonds between the carbonyl oxygens and the charged groups in the side chains (Jarrold 1999; Arteca *et al.* 2001). There would also be a strong driving force to form an α-helix that would be stabilized by strong hydrogen bonds. However, this would not allow 'self solvation' and would generally bring the charges closer together. We will have to develop even more sophisticated mass spectroscopic techniques to gain a good understanding of the denatured state of proteins in a vacuum.

Nanoelectrospray mass spectrometry can be used to estimate the stability of a protein in solution (Benesch *et al.* 2003), but to date there have been no estimates of the stability of a protein in a vacuum. Our guess is that protein stability will be increased more than 10-fold, perhaps to greater than 100 kcal mol^{-1}. It will be interesting to see if we can develop techniques to measure this.

7. IS LIFE POSSIBLE WITHOUT WATER?

The most difficult tasks in living cells are performed by globular proteins in their native states. In a polar solvent such as ethanol, the native state will be unfolded and refolded into α-helices so proteins will lose their biological function. Perhaps this could be avoided if the backbone were less polar. Functional proteins have been constructed in which the backbone is formed by ester bonds instead of amide bonds. These proteins might be more stable in a polar solvent. In a non-polar solvent such as cyclohexane, protein native states will be very stable and perhaps able to perform their biological function. However, both proteins and most substrates will have such low solubility in a non-polar solvent that it is difficult to imagine life as we know it. Perhaps the solubility could be improved if the side chains on the surface of the native protein were all non-polar. Even if this were possible, we would still have to change the solubility of most of the substrates. Evolution has had a long time to develop the living cells we are familiar with. Given time, evolution might be able to come up with a different system in a different solvent that we might call living, but it is hard to imagine what it would be.

REFERENCES

Anderson, D. E., Becktel, W. J. & Dahlquist, F. W. 1990 pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**, 2403–2408.

Arteca, G. A., Reimann, C. T. & Tapia, O. 2001 Proteins *in vacuo*: denaturing and folding mechanisms studied with computer-simulated molecular dynamics. *Mass Spectrom. Rev.* **20**, 402–422.

Benesch, J. L. P., Sobott, F. & Robinson, C. V. 2003 Thermal dissociation of multimeric protein complexes by using nano-electrospray mass spectrometry. *Anal. Chem.* **75**, 2208–2221.

Bolen, D. W. & Baskakov, I. V. 2001 The osmophobic effect; natural selection of a thermodynamic force in protein folding. *J. Mol. Biol.* **310**, 955–963.

Brandts, J. F. 1964 The thermodynamics of protein denaturation. II. A model of reversible denaturation and interpretations regarding the stability of chymotrypsinogen. *J. Am. Chem. Soc.* **86**, 4302–4314.

Chen, J. & Stites, W. E. 2001 Packing is a key selection factor in the evolution of protein hydrophobic cores. *Biochemistry* **40**, 15 280–15 289.

Choy, W. Y., Mulder, F. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D. & Kay, L. E. 2002 Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J. Mol. Biol.* **316**, 101–112.

Cohn, E. J. & Edsall, J. T. 1943 *Proteins, amino acids and peptides.* New York: Hafner Publishing Company.

Cooper, A. 2000 Heat capacity of hydrogen-bonded networks: an alternative view of protein folding thermodynamics. *Biophys. Chem.* **85**, 25–39.

D'Aquino, J. A., Gomez, J., Hilser, V. J., Lee, K. H., Amzel, L. M. & Friere, E. 1996 The magnitude of the backbone conformational entropy change in protein folding. *Proteins Struct. Funct. Genet.* **25**, 143–156.

de Bakker, P. I., Hunenberger, P. H. & McCammon, J. A. 1999 Molecular dynamics simulations of the hyperthermophilic protein sac7d from *Sulfolobus acidocaldarius*: contribution of salt bridges to thermostability. *J. Mol. Biol.* **285**, 1811–1830.

Denisov, V. P., Jonsson, B. H. & Halle, B. 1999 Hydration of denatured and molten globule proteins. *Nature Struct. Biol.* **6**, 253–260.

Dill, K. A. 1990 Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155.

Doig, A. J. & Sternberg, M. J. E. 1995 Side-chain conformational entropy in protein folding. *Protein Sci.* **4**, 2247–2251.

Eisenberg, D. S., Kuriyan, J. & Richards, F. M. 2002 John T. Edsall and advances in protein chemistry. *Adv. Protein Chem.* **62**, xi–xiv.

Elcock, A. H. 1998 The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489–502.

Fauchere, J. L. & Pliska, V. E. 1983 Hydrophobic parameters of amino-acid side-chains from the partitioning of N-acetyl-amino-acid amides. *Eur. J. Med. Chem.* **18**, 369–375.

Fersht, A. R. 1987 The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* **12**, 301–304.

Giletto, A. & Pace, C. N. 1999 Buried, charged, non-ion-paired aspartic acid 76 contributes favorably to the conformational stability of ribonuclease T-1. *Biochemistry* **38**, 13 379–13 384.

Goldenberg, D. P. 2003 Computational simulation of the statistical properties of unfolded proteins. *J. Mol. Biol.* **326**, 1615–1633.

Grimsley, G. R., Shaw, K. L., Fee, L. R., Alston, R. W., Huyghues-Despointes, B. M. P., Thurlkill, R. L., Scholtz, J. M. & Pace, C. N. 1999 Increasing protein stability by altering long-range Coulombic interactions. *Protein Sci.* **8**, 1843–1849.

Guerois, R., Nielsen, J. E. & Serrano, L. 2002 Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **320**, 369–387.

Harpaz, Y., Gerstein, M. & Chothia, C. 1994 Volume changes in protein folding. *Structure* **2**, 641–649.

Hebert, E. J., Giletto, A., Sevcik, J., Urbanikova, L., Wilson, K. S., Dauter, Z. & Pace, C. N. 1998 Contribution of a conserved asparagine to the conformational stability of ribonucleases Sa, Ba and T1. *Biochemistry* **37**, 16 192–16 200.

Hirota-Nakaoka, N. & Goto, Y. 1999 Alcohol-induced denaturation of beta-lactoglobulin: a close correlation to the alcohol-induced alpha-helix formation of melittin. *Bioorg. Med. Chem.* **7**, 67–73.

Honig, B. 1999 Protein folding: from the levinthal paradox to structure prediction. *J. Mol. Biol.* **293**, 283–293.

Huyghues-Despointes, B. M., Thurlkill, R. L., Daily, M. D., Schell, D., Briggs, J. M., Antosiewicz, J. M., Pace, C. N. & Scholtz, J. M. 2003 pK values of histidine residues in ribonuclease Sa: effect of salt and net charge. *J. Mol. Biol.* **325**, 1093–1105.

Jarrold, M. F. 1999 Unfolding, refolding, and the hydration of proteins in the gas phase. *Acc. Chem. Res.* **32**, 360–367.

Kajander, T., Kahn, P. C., Passila, S. H., Cohen, D. C., Lehtio, L., Adolfsen, W., Warwicker, J., Schell, U. & Goldman, A. 2000 Buried charged surface in proteins. *Struct. Fold Des.* **8**, 1203–1214.

Karshikoff, A. & Ladenstein, R. 1998 Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. *Protein Engng* **11**, 867–872.

Kauzmann, W. 1959 Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* **14**, 1–63.

Khan, F., Chuang, J. I., Gianni, S. & Fersht, A. R. 2003 The kinetic pathway of folding of barnase. *J. Mol. Biol.* **333**, 169–186.

Klapper, M. H. 1971 On the nature of the protein interior. *Biochim. Biophys. Acta* **229**, 557–566.

Klibanov, A. M. 2001 Improving enzymes by using them in organic solvents. *Nature* **409**, 241–246.

Laurents, D. L., Perez-Canadillas, J. M., Santoro, J., Rico, M., Schell, D., Pace, C. N. & Bruix, M. 2001 Solution structure and dynamics of ribonuclease Sa. *Proteins Struct. Funct. Genet.* **44**, 200–211.

Laurents, D. V. (and 13 others) 2003 Charge–charge interactions are key determinants of the pK values of ionizable groups in ribonuclease Sa (pI = 3.5) and a basic variant (pI = 10.2). *J. Mol. Biol.* **325**, 1077–1092.

Lesser, G. J. & Rose, G. D. 1990 Hydrophobicity of amino acid subgroups in proteins. *Proteins Struct. Funct. Genet.* **8**, 6–13.

Liu, Y. & Bolen, D. W. 1995 The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry* **34**, 12 884–12 891.

Loladze, V. V., Ermolenko, D. N. & Makhatadze, G. I. 2002 Thermodynamic consequences of burial of polar and nonpolar amino acid residues in the protein interior. *J. Mol. Biol.* **320**, 343–357.

Lomize, A. L., Reibarkh, M. Y. & Pogozheva, I. D. 2002 Interatomic potentials and solvation parameters from protein engineering data for buried residues. *Protein Sci.* **11**, 1984–2000.

Luo, P. & Baldwin, R. L. 1997 Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36**, 8413–8421.

Malissard, M. & Berger, E. G. 2001 Improving solubility of catalytic domain of human beta-1,4-galactosyltransferase 1 through rationally designed amino acid replacements. *Eur. J. Biochem.* **268**, 4352–4358.

Marti, D. N. & Bosshard, H. R. 2003 Electrostatic interactions in leucine zippers: thermodynamic analysis of the contributions of Glu and His residues and the effect of mutating salt bridges. *J. Mol. Biol.* **330**, 621–637.

Matthews, B. W. 1995 Studies on protein stability with T4 lysozyme. *Adv. Protein Chem.* **46**, 240–311.

Mayor, U., Grossmann, J. G., Foster, N. W., Freund, S. M. & Fersht, A. R. 2003 The denatured state of engrailed homeodomain under denaturing and native conditions. *J. Mol. Biol.* **333**, 977–991.

Miller, W. G. & Goebel, C. V. 1968 Dimensions of protein random coils. *Biochemistry* **7**, 3925–3935.

Millet, I. S., Townsley, L. E., Chiti, F., Doniach, S. & Plaxco, K. W. 2002 Equilibrium collapse and the kinetic 'foldability' of proteins. *Biochemistry* **41**, 321–325.

Mirsky, A. E. & Pauling, L. 1936 On the structure of native, denatured, and coagulated proteins. *Proc. Natl Acad. Sci. USA* **22**, 439–447.

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

Nozaki, Y. & Tanford, C. 1971 The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions; establishment of a hydrophobicity scale. *J. Biol. Chem.* **246**, 2211–2217.

Pace, C. N. 1975 The stability of globular proteins. *CRC Crit. Rev. Biochem.* **3**, 1–43.

Pace, C. N. 1995 Evaluating contribution of hydrogen bonding and hydrophobic bonding to protein folding. *Meth. Enzymol.* **259**, 538–554.

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

Pace, C. N. 2001 Polar group burial contributes more to protein stability than nonpolar group burial. *Biochemistry* **40**, 310–313.

Pace, C. N. & Scholtz, J. M. 1997 Measuring the conformational stability of a protein. In *Protein structure: a practical approach* (ed. T. E. Creighton), pp. 299–321. Oxford: IRL Press.

Pace, C. N., Heinemann, U., Hahn, U. & Saenger, W. 1991 Ribonuclease T1: structure, function and stability. *Angew. Chem. Int. Ed. Engl.* **30**, 343–360.

Pace, C. N. (and 10 others) 1998 Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3. *J. Mol. Biol.* **279**, 271–286.

Pace, C. N., Alston, R. W. & Shaw, K. L. 2000 Charge–charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci.* **9**, 1395–1398.

Pace, C. N., Horn, G., Hebert, E. J., Bechert, J., Shaw, K., Urbanikova, L., Scholtz, J. M. & Sevcik, J. 2001 Tyrosine hydrogen bonds make a large contribution to protein stability. *J. Mol. Biol.* **312**, 393–404.

Pauling, L., Corey, R. B. & Branson, H. R. 1951 The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl Acad. Sci. USA* **37**, 205–211.

Privalov, P. L. & Makhatadze, G. I. 1993 Contribution of hydration to protein folding thermodynamics II. Entropy and Gibbs energy of hydration. *J. Mol. Biol.* **232**, 660–679.

Qu, Y., Bolen, C. L. & Bolen, D. W. 1998 Osmolyte-driven contraction of a random coil protein. *Proc. Natl Acad. Sci. USA* **95**, 9268–9273.

Radzicka, A. & Wolfenden, R. 1988 Comparing the polarites of the amino acids; side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. *Biochemistry* **27**, 1644–1670.

Schmitke, J. L., Stern, L. J. & Klibanov, A. M. 1997 The crystal structure of subtilisin Carlsberg in anhydrous dioxane and its comparison with those in water and acetonitrile. *Proc. Natl Acad. Sci. USA* **94**, 4250–4255.

Sevcik, J., Dauter, Z., Lamzin, V. S. & Wilson, K. S. 1996 Ribonuclease from *Streptomyces aureofaciens* at atomic resolution. *Acta Crystall.* D **52**, 257–271.

Sevcik, J., Lamzin, V. S., Dauter, Z. & Wilson, K. S. 2002 Atomic resolution data reveal flexibility in the structure of RNase Sa. *Acta Crystall.* D **58**, 1307–1313.

Shaw, K. L., Grimsley, G. R., Yakovlev, G. I., Makarov, A. A. & Pace, C. N. 2001 The effect of net charge on the solubility, activity, and stability of ribonuclease Sa. *Protein Sci.* **10**, 1206–1215.

Shirley, B. A., Stanssens, P., Hahn, U. & Pace, C. N. 1992 Contribution of hydrogen bonding to the conformational stability of ribonuclease T1. *Biochemistry* **31**, 725–732.

Shortle, D. 2002 The expanded denatured state: an ensemble of conformations trapped in a locally encoded topological space. *Adv. Protein Chem.* **62**, 1–23.

Spolar, R. S. & Record, M. T. J. 1994 Coupling of local folding to site-specific binding of proteins to DNA. *Science* **263**, 777–784.

Steyaert, J. 1997 A decade of protein engineering on ribonuclease T1. *Eur. J. Biochem.* **247**, 1–11.

Stickle, D. F., Presta, L. G., Dill, K. A. & Rose, G. D. 1992 Hydrogen bonding in globular proteins. *J. Mol. Biol.* **226**, 1143–1159.

Takano, K., Scholtz, J. M., Sacchettini, J. C. & Pace, C. N. 2003 The contribution of polar group burial to protein stability is strongly context-dependent. *J. Biol. Chem.* **278**, 31 790–31 795.

Tanaka, S., Oda, Y., Ataka, M., Onuma, K., Fujiwara, S. & Yonezawa, Y. 2001 Denaturation and aggregation of hen egg lysozyme in aqueous ethanol solution studied by dynamic light scattering. *Biopolymers* **59**, 370–379.

Tanford, C. 1962 Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. *J. Am. Chem. Soc.* **84**, 4240–4247.

Tanford, C. 1968 Protein denaturation. *Adv. Protein Chem.* **23**, 121–282.

Tsai, J., Taylor, R., Chothia, C. & Gerstein, M. 1999 The packing density in proteins; standard radii and volumes. *J. Mol. Biol.* **290**, 253–266.

Uversky, V. N. 2002 Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* **11**, 739–756.

Wolynes, P. G. 1995 Biomolecular folding *in vacuo*!!!(?) *Proc. Natl Acad. Sci. USA* **92**, 2426–2427.

Xiao, L. & Honig, B. 1999 Electrostatic contributions to the stability of hyperthermophilic proteins. *J. Mol. Biol.* **289**, 1435–1444.

Discussion

F. Franks (*BioUpdate Foundation, London, UK*). You state rightly that denatured state(s) of proteins are much less stable than the native configurations. Is not the reason that denaturants (chaotropes) rob water of one of its most important properties: the ability to repel apolar residues, so-called hydrophobic hydration which in turn gives rise to the compact folded structures? Once this ability has been removed, water behaves just like any other polar solvent, and the protein behaves almost like any other polyelectrolyte, with no driving force for specific folding. Is this not borne out by the observation that, for stable folded states to exist, the peptide chain requires an apolar residue content of *ca.* 50% and that these residues tend to be those most completely conserved, e.g. the globins, and are identical in structures if not in functions for all species?

C. N. Pace. I agree with your assessment of hydrophobic groups and denaturants, but I will put it a different way. The hydrophobic parts of a protein are more soluble in the presence of the denaturant than they are in water and in part this is why proteins unfold when a denaturant is added. However, it turns out that all of the constituent groups of a protein, charged, polar and non-polar, are more soluble in the presence of denaturant than in water so they all contribute to the unfolding by denaturants. So, in water, the hydrophobic effect contributes to protein stability, but we think the burial of the peptide groups may be equally important. In a vacuum or an organic solvent, we think it is the hydrophilic effect, i.e. the burial of peptide groups that makes the major contribution to protein stability.

F. Franks. Can you explain the observation that biopolymers in which hydrophobia cannot play a role, e.g. many polysaccharides, also form stable ordered structures in aqueous solution, and that these structures can be destabilized by changes in the aqueous solvent medium?

C. N. Pace. My guess would be that the intramolecular hydrogen bonds and van der Waals interactions formed by groups in the biopolymer are stronger than the intermolecular interactions that the same groups would form with water when the structure is disrupted. Changes in the aqueous solvent medium could then shift the equilibrium in either direction.

F. Franks. In any discussion of protein 'in vacuum', one can ask: how are such proteins synthesized in the absence of water, because once they have been synthesized, they cannot be *completely* dehydrated without losing their native structures? As has been repeatedly mentioned in this meeting, some water molecules are always integral to the polypeptide and are required to turn it into a 'protein'.

C. N. Pace. I used to have a similar view: some bound molecules of water cannot be removed no matter how long you freeze-dry a protein, and bad things happen to some proteins, e.g. barnase, when they are freeze-dried. However, the people doing mass spectroscopy on proteins should know with certainty if a water molecule is bound to their proteins, and, in some cases, none is bound, yet the protein is thought to be folded. This idea is still controversial. See the article by Jarrold (1999) for a discussion of the hydration of proteins in a vacuum.

K. Wilson (*Department of Chemistry, University of York, York, UK*). All weak forces must be simultaneously studied in a protein fold.

C. N. Pace. I think Professor Wilson is referring to this possible problem: the various forces are large, over 100 kcal mol^{-1}, so how can you learn about the small contributions of less than 2 kcal mol^{-1} observed for most mutations? The thinking is that when an −OH group is removed, for example, it is possible that a ripple effect

occurs so that there are small changes in the forces over most of the molecule and this is what gives rise to changes in stability rather than just local effects. We now have high-resolution structures of hundreds of mutant proteins. In most cases, only small structural changes are observed and these are generally localized near the site of the mutation (Matthews 1995). In an analysis of a large number of mutants with known crystal structures, conclusions similar to those reported in our paper are reached (Lomize *et al.* 2002). Occasionally something more interesting occurs. In a case Professor Wilson worked on (Hebert *et al.* 1998), we observed a significant rearrangement of a surface loop in an Asn 39 to Ser mutant of ribonuclease Sa.

R. M. Daniel (*Department of Biological Sciences, University of Waikato, Hamilton, New Zealand*). Comment: Mullaney (1966) looked at protein stability *in vacuo*, and saw high stability at temperatures greater than 150 °C.

C. N. Pace. The remarkable paper by Mullaney (1966) indicates that both ribonuclease and trypsin would have half lives of over 5 min at temperatures above 200 °C and a pressure of 40 mm Hg. This is consistent with our thought that proteins might be much more stable in a vacuum than they are in water.

P. J. Halling (*Department of Chemistry, University of Strathclyde, Glasgow, UK*). I think you are right that native protein structures become more stable when water is replaced by a non-polar medium, as I have believed for some time. My question relates to your mutation experiments changing Tyr to Phe. When the Tyr was hydrogen bonded, the mutation leaves a hydrogen bond partner now unsatisfied. How do you allow for this contribution to the observed energy change? That explains what happens structurally, but what about the energetics? If that water is now not forming an optimal hydrogen bond pattern, will there not be an energetic penalty for this?

C. N. Pace. This is an important point, and the possibilities were discussed in our first paper on hydrogen bonding (Shirley *et al.* 1992) and more recently (Pace 1995). We now have crystal structures available for many hydrogen-bonding mutants, and they show that the hydrogen bond partner is generally hydrogen bonded to a water molecule in the mutant (Pace *et al.* 2001). Thus, the remaining hydrogen bond partner should make only a small contribution to the change in stability since it is hydrogen bonded to water in both the folded and unfolded states. Consequently, with regard to hydrogen bonds, the main contribution to the change in stability is the one or more intramolecular hydrogen bonds that are lost. However, our more recent studies (Takano *et al.* 2003) suggest that van der Waals interactions and electrostatic interactions at longer range than hydrogen bonds may make a more important contribution than the hydrogen bonds themselves. Thus, the sites where polar groups are buried in folded proteins have been carefully selected to optimize the interaction of the polar group with the other groups in the molecule. So when one member of the hydrogen bond pair is removed, you lose more than just the intramolecular hydrogen bonds.

D. S. Clark (*Department of Chemical Engineering, University of California at Berkeley, Berkeley, CA, USA*). You mentioned that you expect proteins to be less stable when they are more rigid. Why do you expect this to be the case?

C. N. Pace. There is much freer rotation around the bonds in the unfolded states of a protein than in the folded states. Consequently, the main force favouring the unfolding of a protein is conformational entropy. If the molecules in the folded state are more rigid, I think the entropy change favouring unfolding would be greater than when they are less rigid. If the increased rigidity increased the van der Waals interactions in the interior of a protein, it might lead to increased stability, but it is hard to imagine that the protein interior could be more tightly packed than it is for most proteins.

A. Purkiss (*School of Crystallography, Birkbeck College, London, UK*). When considering the effect of charge in protein solubility, what is the importance of partial charges, such as in amide bonds on the solubility of the protein?

C. N. Pace. My guess is that the more exposed amide groups on the surface of a protein there are, the greater is the solubility. However, I am not sure that this has been studied. We have begun a project to gain a better understanding of what determines the solubility of folded and unfolded proteins. Some of our preliminary results are shown in figure 4 and table 5.

Additional reference

Mullaney 1966 Dry thermal inactivation of trypsin and ribonuclease. *Nature* **210**, 953.

GLOSSARY

GdnHCl: guanidine hydrochloride
NMR: nuclear magnetic resonance
RNase Sa: ribonuclease Sa