Downloaded from rsta.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Protein Stability: Experimental Data from Protein Engineering

Alan R. Fersht, Sophie E. Jackson and Luis Serrano

Phil. Trans. R. Soc. Lond. A 1993 **345**, 141-151
doi: 10.1098/rsta.1993.0125

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. Lond. A* go to: **http://rsta.royalsocietypublishing.org/subscriptions**

This journal is © 1993 The Royal Society

Downloaded from rsta.royalsocietypublishing.org

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

Protein stability: experimental data from protein engineering

By Alan R. Fersht, Sophie E. Jackson and Luis Serrano
Cambridge Centre for Protein Engineering, MRC Centre, Hills Road, Cambridge CB2 2QH, U.K. and the Department of Chemistry, Lensfield Road, Cambridge, U.K.

All of molecular recognition, from the binding of substrates by enzymes, information transfer in replicating and processing the genetic information to the folding of proteins, is dominated by non-covalent interactions. Perhaps the most difficult challenge is understanding protein folding because each group in the molecule has to recognize with which ones it has to pair. Protein engineering is providing an experimental entry to determine the magnitude, nature and importance of the various levels of recognition in protein folding. In addition to providing the energetics of specific interactions, fundamental information has been given on the energetics of burial of hydrophobic and hydrophilic solvent-accessible surface areas and their specific roles in stabilizing protein cores and helices.

1. Introduction

Molecular recognition in biology is dominated by the chemistry of the non-covalent bond. Our lack of knowledge of the precise values for the energetics of these interactions is most apparent when analysing the stability of proteins and the energetics of enzyme catalysis. Computer methods are currently inadequate for the accurate calculation of the necessary energetics (Presnell *et al*. 1992; Van Gunsteren & Mark 1992). The free energy of unfolding of small proteins is often between 20 and 60 kJ mol^{-1}. This small energy balance is the difference between two very large numbers, the free energies of all the non-covalent interactions in the folded and unfolded states, both of which amount to thousands of kJ mol^{-1} and neither of which can be calculated with high precision. It has been possible for the past decade, however, to acquire basic data on the strengths of the common interactions from experiments based on protein engineering. Groups that are involved in clearly defined interactions within a structure or in a complex are changed in a chemically-sensible manner. The change in stability on mutation is then measured to give an empirical value for the interaction. These data are accumulated to give an experimental data base for either direct use or for acting as benchmarks for refining theoretical calculation. The results from protein engineering have been discussed in depth recently elsewhere (Fersht & Serrano 1993). In this paper, we shall concentrate on some selected topics that are general themes that have been the subject of much discussion in the meeting.

It is important first to define what is actually measured from the protein engineering experiments. Suppose we mutate one side chain in a protein to a smaller one and so delete a group that makes a particular interaction. The change in the free

Phil. Trans. R. Soc. Lond. A (1993) **345**, 141–151
Printed in Great Britain
© 1993 The Royal Society

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to *Philosophical Transactions: Physical Sciences and Engineering.*
www.jstor.org

Downloaded from rsta.royalsocietypublishing.org

energy of unfolding of the protein on mutation does not give directly the binding energy of the group that was deleted but gives instead the relative energies of having the wild-type and mutated side chain. We thus measure an effect on specificity rather than a direct binding energy (Fersht 1987, 1988). But this is precisely the value that is required for use in protein design and for calculation of the specificity of interactions. Further, the measurements are analogous to those in many model studies, and so provide direct comparisons. For example, the importance of the hydrophobic effect has been measured from the free energies of transfer of the side chains of amino acids from water to organic solvents, relative to the transfer of the side chain of glycine. Similarly, we measure the free energy of transfer to the hydrophobic core of a protein of the side chains of amino acid residues relative to the transfer of a reference side chain.

2. The hydrophobic effect

The hydrophobic effect was well-studied before the introduction of protein engineering methods because it is readily studied in enzyme–ligand interactions (reviewed by Fersht 1985, ch. 11). The effect was established to be variable in protein–ligand interactions and often far higher than in model systems because of a combination of the effects of dispersion energies and the existence of preformed cavities in enzymes, either empty or filled with water. Model experiments (Chothia 1976; Eisenberg & McLachlan 1986; Ooi *et al.* 1987) find a linear relation between the free energy of transfer of a small solute to a hydrophobic solvent and the burial of solvent-accessible hydrophobic surface area. The constant of proportionality was found in the earliest studies to be 80–100 J mol^{-1} $Å^{-2}$ ($\equiv$ 2.5–3.3 kJ mol^{-1} per methylene group).

The importance of the hydrophobic core is most easily probed by using mutations that create small deletions in side chains without changing the geometry (Kellis *et al.* 1988). Experiments on barnase (Kellis *et al.* 1988, 1989; Serrano *et al.* 1992*a*), staphylococcal nuclease (Shortle *et al.* 1990), gene5 of bacteriophage f1 (Sandberg & Terwilliger 1991) and T4 lysozyme (Eriksson *et al.* 1992) find that the truncation of the side chains of Ile to Val, Ala and Gly, Leu to Ala and Gly, Val to Ala and Gly and so on lose up to 10 kJ mol^{-1} per methylene group deleted, and typically 6 ± 2 kJ mol^{-1} ($\pm$s.d.) (Yutani *et al.* 1987; Kellis *et al.* 1988, 1989; Matsumura *et al.* 1988; Daopin *et al.* 1991; Shortle *et al.* 1990; Serrano *et al.* 1992*a*; Eriksson *et al.* 1992). These changes are much smaller, however, than observed for the binding of hydrophobic side chains of amino acids to aminoacyl-tRNA synthetases that have evolved to bind them as tightly as possible. These bind each methylene group by some 13–14.5 kJ mol^{-1}, the isoleucyl side chain being bound more tightly than that of alanine by 40 kJ mol^{-1} (Fersht 1985).

Rearrangement of structure on mutation must always compensate for the loss of missing interactions (Kellis *et al.* 1989). Eriksson *et al.* (1992) have solved the crystal structures of mutants of T4 lysozyme which have had deletion mutations in the hydrophobic core. The maximum changes in energy occurs where the mutant has the full-sized empty cavity. As the cavity collapses due to neighbouring side chains moving in, there are correspondingly smaller changes in energy. As hydrophobic interactions are cumulative, the effects of removing several methylene groups causes the largest decreases in stability, for example Ile or Leu mutated to Ala can cost 20–30 kJ mol^{-1} (Serrano *et al.* 1992*a*). The hydrophobic core is more tolerant to

Downloaded from rsta.royalsocietypublishing.org

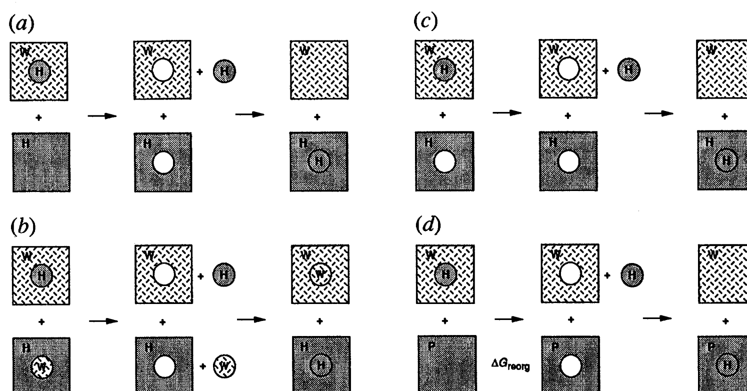

Figure 1. Illustration of notional steps in the transfer of hydrophobic groups from water (W) to hydrocarbon (H) solvents: (*a*) to hydrocarbon; (*b*) to hydrocarbon (= protein) with a preformed cavity that is filled with water; (*c*) to hydrocarbon with a preformed empty cavity; and (*d*) to a cavity in a protein that collapses when the hydrocarbon is not present in it.

substitution of hydrophobic side chains by others of different stereochemistry than would be expected (Lim *et al.* 1992; Hurley *et al.* 1992).

There is a fundamental difference, discussed in the next section, between the hydrophobic effect as measured from simple model partition experiments and those involving proteins. The transfer of a hydrophobic solute from water to organic solvent can be divided into the following notional steps (figure 1): (i) transfer of the hydrophobic solute from water to vacuum; (ii) collapse of the cavity in water; (iii) opening a cavity in the hydrophobic phase; and (iv) transfer of solute to the cavity. The transfer of a hydrophobic solute to a pre-existing cavity in a protein does not have step (iii) in the former process and is thus more favourable. If the cavity is empty, then the transfer will be more favourable for transfer to the protein by the energy expended in opening a cavity. If the cavity is filled with water then it will be slightly more favourable still. A macroscopic surface tension model, which follows next, predicts that, all things being equal, the transfer to a water-filled cavity is exactly twice as energetically favourable as transfer to an organic solvent (Kellis *et al.* 1989).

3. Simplified macroscopic model for comparing the free energies of transfer of side chains from water to hydrocarbons with transfer to proteins

The thermodynamic analysis of the burial of a side chain of an amino acid in the folding of a protein is analogous to that for the transfer of the side chain of that amino acid from water to a hydrocarbon or other solvent. There is an identical standard state for both processes: the standard state for the side chain in a fully-unfolded denatured state in water for analysing the process of protein folding is identical to that for a model peptide in water for analysing the process of transfer to an organic solvent. Thus, any factors that influence the free energy of the side chain in water are identical for both processes and so will cancel out when comparing the two.

A scheme for analysing the transfer processes is given in figure 1. The following analysis is for a simple classical model of surface tension. It is not correct in detail but it does illustrate some physical principles in comparing protein folding and model

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

experiments on the transfer of solutes from water to hydrocarbons. The key assumption for simplicity is that on comparing the transfer of a hydrophobic side chain of an amino acid from water to a hydrocarbon solvent with the transfer of the same hydrophobic side chain to a buried hydrophobic region in a protein, the hydrophobic region has identical characteristics to the hydrocarbon solvent. (In practice, the effects are variable in a protein because of specific interactions, variations in dispersion energies and structural changes on mutation.)

According to the Dupré (1869) equation, the 'work of adhesion', w, between two immiscible liquids is given by

$$w = \gamma_1 + \gamma_2 - \gamma_{1/2}, \tag{1}$$

where γ_1 is the surface tension of liquid 1, γ_2 is that of liquid 2 and $\gamma_{1/2}$ is the interfacial surface tension of the two liquids. Equation (1) may be applied to the examples in figure 1 by dividing up the transfer of a hydrophobic solute H from water to a hydrocarbon or to an equivalent hydrophobic region in a protein into notional steps. For a typical hydrocarbon $\gamma_{\mathrm{W}} = 435\ \mathrm{J\ mol^{-1}\ \AA^{-2}}$, $\gamma_{\mathrm{H}} = 108\ \mathrm{J\ mol^{-1}\ \AA^{-2}}$, $\gamma_{\mathrm{H/W}} = 304\ \mathrm{J\ mol^{-1}\ \AA^{-2}}$ (calculated from the surface tension data for water and hexane (Hildebrand 1979)).

Case A. Transfer of hydrocarbon H from water to hydrocarbon.

The transfer of a hydrocarbon solute of surface area A from water to hydrocarbon is considered to be composed of the processes: removal to vacuum of hydrocarbon, which removes the hydrocarbon/water interface ($w = -\gamma_{\mathrm{H/W}}$), and creates a hydrocarbon/vacuum and a water/vacuum interface ($w = \gamma_{\mathrm{H}} + \gamma_{\mathrm{W}}$); formation of a cavity in the hydrocarbon that creates a hydrocarbon/vacuum interface ($w = \gamma_{\mathrm{H}}$); transfer of the hydrocarbon solute to the hydrocarbon and collapse of the water/vacuum cavity, which loses two hydrocarbon/vacuum surfaces and a water/vacuum surface ($w = -2\gamma_{\mathrm{H}} - \gamma_{\mathrm{W}}$). The free energy of transfer, $\Delta G_{\mathrm{trans}}$, is given by $A\Sigma w$; i.e.

$$\Delta G_{\mathrm{trans}} = A(-\gamma_{\mathrm{H/W}} + \gamma_{\mathrm{H}} + \gamma_{\mathrm{W}} + \gamma_{\mathrm{H}} - 2\gamma_{\mathrm{H}} - \gamma_{\mathrm{W}}) \tag{2}$$

$$= -A\gamma_{\mathrm{H/W}}. \tag{3}$$

This may be seen directly by comparing the right-hand side of figure 1*a* with the left, which shows that the net process is the loss of one water/hydrocarbon interface.

Case B. Transfer of hydrocarbon H from water to hydrocarbon (protein) with preformed cavity that is filled with water.

$$\Delta G_{\mathrm{trans}} = A(-2\gamma_{\mathrm{H/W}} + 2\gamma_{\mathrm{H}} + 2\gamma_{\mathrm{W}} - 2\gamma_{\mathrm{W}} - 2\gamma_{\mathrm{H}}) \tag{4}$$

$$= -2A\gamma_{\mathrm{H/W}}. \tag{5}$$

That is, the free energy of transfer is *twice* that calculated for case A. This is expected from comparing the right-hand side of figure 1*b* with the left, which shows that the net process in transfer is the loss of two water/hydrocarbon interfaces.

Case C. Transfer of hydrocarbon H from water to hydrocarbon (protein) which has preformed cavity that is empty.

This is the same as case A, apart from a term γ_{H} that arises from forming the cavity in the hydrocarbon

$$\Delta G_{\mathrm{trans}} = A(-\gamma_{\mathrm{H/W}} + \gamma_{\mathrm{W}} + \gamma_{\mathrm{H}} - 2\gamma_{\mathrm{H}} - \gamma_{\mathrm{W}}) \tag{6}$$

$$= -A(\gamma_{\mathrm{H/W}} + \gamma_{\mathrm{H}}). \tag{7}$$

The free energy change is greater than that for case A but less than that for case B ($\gamma_{\mathrm{H}} = 108\ \mathrm{J\ mol^{-1}\ \AA^{-2}}$ and $\gamma_{\mathrm{H/W}} = 304\ \mathrm{J\ mol^{-1}\ \AA^{-2}}$ for hexane and water).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

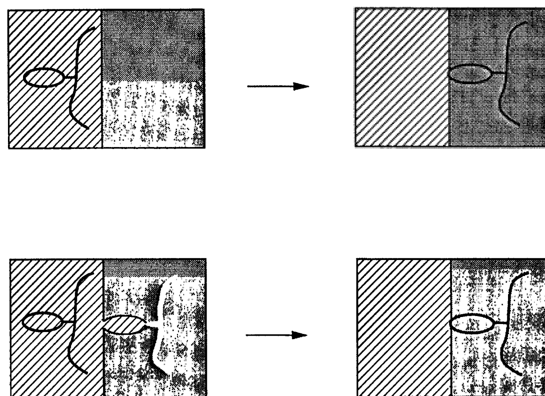

Figure 2. Cartoon comparing the transfer of a hydrophobic group side chain to a hydrocarbon and to a protein with a preformed cavity on the surface.

Case D. Transfer of hydrocarbon H from water to mutant protein within which the cavity has collapsed.

Consider a mutant enzyme in which a hydrophobic side chain has been deleted to leave a cavity but the cavity collapses. As collapse is a spontaneous process, it is accompanied by a favourable energy change ΔG_{reorg}, the reorganization energy of the protein, that has an algebraically negative sign. The non-covalent energy term for the transfer of the deleted hydrophobic side chain into its correct site in the mutant has all the energy terms described for case A plus ΔG_{reorg} (the notional step of formation of the cavity in the mutant enzyme creates a hydrocarbon/vacuum interface and costs also the energy of reorganizing the surrounding side chains back to their positions in the wild-type enzyme ($w = \gamma_{\text{H}} + \Delta G_{\text{reorg}}$)).

$$\Delta G_{\text{trans}} = A(-\gamma_{\text{H/W}} + \gamma_{\text{H}} + \Delta G_{\text{reorg}} + \gamma_{\text{W}} + \gamma_{\text{H}} - 2\gamma_{\text{H}} - \gamma_{\text{W}}) \tag{8}$$

$$\Delta G_{\text{trans}} = -A(\gamma_{\text{H/W}} - \Delta G_{\text{reorg}}). \tag{9}$$

The free energy of transfer is less than that of case A.

Figure 2 illustrates the transfer of a side chain to the surface of a protein. The analysis is identical to that of case B. $\Delta G_{\text{trans}} = -2A\gamma_{\text{H/W}}$. It has been found – see later – that ΔG_{trans} for Ala versus Gly to the surfaces of the helices in barnase (160–220 J mol^{-1} Å^{-2}) is twice that for model experiments for the hydrophobic side chains to hydrocarbons (80–100 J mol^{-1} Å^{-2}) (Serrano *et al.* 1992*a*). Although the relative values are consistent, the absolute values are less than the value of 304 J mol^{-1} Å^{-2} expected from the macroscopic constants for hexane and water. There have been attempts to reconcile the macroscopic values with the experimental data (Nicholls *et al.* 1991; Sharp *et al.* 1991*a*, *b*).

The above analysis shows that, in principle: (i) the transfer of an amino acid side chain to a water filled buried or surface cavity should be twice as favourable as transfer to a hydrocarbon solvent of identical characteristics; (ii) the energy of transfer to an empty cavity is intermediate between that to hydrocarbon solvent and a water-filled cavity; and (iii) the energy of transfer to a collapsed cavity is less than that to a hydrocarbon solvent. In practice, energies have been found in protein engineering studies that are higher than the above, and even higher values have been found in enzyme–ligand interactions.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

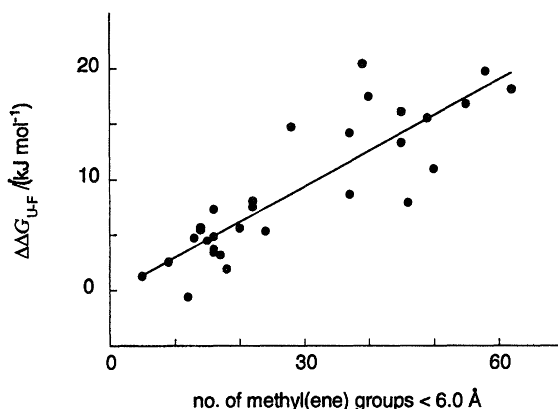

Figure 3. Change in free energy of unfolding on the mutation of Ile or Val to Val or Ala versus the number of methylene groups within a sphere of radius 6 Å of the methylene groups that are deleted. The data are from barnase (Serrano *et al.* 1992*a*) and the barley CI-2 inhibitor (S. E. Jackson, N. elMasry, M. Moracci, C. Johnson & A. R. Fersht, unpublished data).

Importance of packing

A second difference between the results from models and with proteins results from the packing density of atoms in proteins. This is illustrated experimentally from a correlation between the observed change in energy on deletion of methylene groups from hydrophobic side chains in a core and the number of $-CH_2-$ and CH_3- groups in a sphere of radius 6 Å surrounding the methylene group that is deleted (Serrano *et al.* 1992*a*) and S. E. Jackson, N. elMasry, M. Moracci, C. Johnson & A. R. Fersht (unpublished data). It seems that the higher the packing density, the stronger the van der Waals' interactions with the target methylene group and also amongst the surrounding side chains so that they are less likely to collapse into the cavity formed on mutation (figure 3).

4. Hydrogen bonding

The properties of hydrogen bonds in proteins are quite variable (Baker & Hubbard 1984). Mutation of a large number of hydrogen bonds in the tyrosyl-tRNA synthetase found that removal of the partner of an uncharged donor/acceptor lowers stability by 2–8 kJ mol^{-1} and removal of the partner of a charged group by 12–20 kJ mol^{-1} (Fersht *et al.* 1985). The range of 2–8 kJ mol^{-1} is also that for the net stabilization energy of the hydrogen bond (Fersht 1987). The results of a large number of mutations in proteins confirm that these ranges hold also for stability of proteins (Shirley *et al.* 1992; Serrano *et al.* 1992*a*). For hydrogen bonds in which there is no access of water, removal of the uncharged partner/donor results in a decrease in stability of 4–8 kJ mol^{-1} per hydrogen bond. On the other hand in those cases in which there is access of water, the decrease in stability is smaller (0–2 kJ mol^{-1}).

5. The stability of α-helices

Some of the most general information about the stability of particular elements of structure gathered so far from protein engineering experiments concerns the α-helix. About one-third of amino acid residues in proteins are in α-helices. The first and last four residues cannot make the intrahelical hydrogen bonds between the backbone CO groups of one turn and the NH groups of residues in the next. The ends of helices thus

Downloaded from rsta.royalsocietypublishing.org

make hydrogen bonds with either solvent or groups in the protein (Presta & Rose 1988; Richardson & Richardson 1988). The first and last residue of a helix are called the N- and C-caps (Richardson & Richardson 1988).

Helix forming propensities of amino acids

(*a*) *Glycine versus alanine*

The relative merits of Gly and Ala for stabilizing a helix is both positional and context dependent (Serrano *et al.* 1992*b*, *c*). Three principles are involved. First, Gly is inherently more destabilizing than Ala because Gly has more conformational freedom in the unfolded state (Matthews *et al.* 1987). Second, Ala buries more solvent-accessible hydrophobic surface area on folding from an extended conformation than does Gly (Richards & Richmond 1978). The amount buried depends on the nature of the surrounding side chains. An empirical equation was found (Serrano *et al.* 1992*b*, *c*) for the relative effects of Ala and Gly on stability ($\Delta\Delta G_{\mathrm{Gly}\rightarrow\mathrm{Ala}}$) and the difference in the solvent-exposed hydrophobic surface area of the helix containing Ala minus that containing Gly(ΔA_{HP}) in internal solvent-exposed positions:

$$\Delta\Delta G_{\mathrm{Gly}\rightarrow\mathrm{Ala}} = 7.7 - 0.23\Delta A_{\mathrm{HP}} \quad (\mathrm{kJ\ mol^{-1}}). \tag{10}$$

Third, the side chain of Ala can block the solvation of NH and CO groups that need to make hydrogen bonds with solvent (Serrano & Fersht 1989). This is important for the C- and N-termini of the helices. An empirical correlation was found (Serrano *et al.* 1992) that relates $\Delta\Delta G_{\mathrm{Gly}\rightarrow\mathrm{Ala}}$ to both ΔA_{HP} and difference in solvent-accessible surface area of NH and CO groups in the helix that require solvation (ΔA_{HB} = area of Ala-containing helix − area of Gly-containing helix):

$$\Delta\Delta G_{\mathrm{Gly}\rightarrow\mathrm{Ala}} = 6.7 - 0.19\Delta A_{\mathrm{HP}} + 0.79\Delta A_{\mathrm{HB}} \quad (\mathrm{kJ\ mol^{-1}}). \tag{11}$$

The relative free energies of Ala and Gly calculated for a model polyalanine helix shows that Ala is more stable in the middle positions of an α-helix but Gly is preferred at the ends (Serrano *et al.* 1992*b*, *c*). This is more noticeable at the N-terminus since the side chain points backwards along a helix.

(*b*) *Residues at C-caps*

The rank order of preference at the two C-caps of barnase is: Gly ≫ Arg, His > Asn, Lys > Ala, Ser, Thr > Asp > Val. The His at the C-cap of helix$_1$ and the Gly at the C-cap of helix$_2$ in barnase are in the α_{L} conformation with positive ϕ and ψ angles. This is a common motif of helices: some 30% of helices have their C-cap in such a conformation and 60% of these are a Gly (Preissner & Bork 1991). This can readily be rationalized by the empirical relationship (11) between stabilization energy and the degree of solvent exposure of groups that require hydrogen bonding with solvent. An α_{L} conformation at the C-terminus of a model polyalanine helix exposes 32 Å^2 of solvent-accessible hydrophilic surface area and buries 26 Å^2 more hydrophobic surface area than when in the regular α_{R} conformation. This could contribute 13 kJ mol^{-1} of energy to compensate for the unfavourable conformation. Gly is the most favourable residue at the C-cap of a helix because the lack of a side chain allows it to adopt an α_{L} conformation without strain and it also allows the greatest solvent exposure of the C-terminal of a helix. Amino acid residues that can make hydrogen bonds with the exposed CO groups are stabilizing with respect to Ala. The alignment of dipoles of the polypeptide backbond parallel to the axis of an α-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

helix cause a net dipole moment with its positive pole at the N-terminus and negative pole at the C-terminus (Hol 1987). Positively charged residues can interact favourably with negative pole at the C-terminus.

(c) *Residues at* N-*caps*

The rank order of preference at the N-caps of barnase is: Asp, Thr, Ser > Asn, Gly > Glu, Gln > His > Ala > Val. Gly allows the greatest exposure of the unpaired NH groups to solvent, bulky hydrophobics inhibit solvation, hydrogen bond acceptors are particularly effective at stabilizing the NH groups by forming intramolecular hydrogen bonds (mainly to residue N-cap + 3) and negative charges can interact favourably with the helix dipole.

(d) *Helix–dipole charge interactions*

The negatively charged carboxylate of Asp as a hydrogen bond acceptor at the N-cap stabilizes barnase and T4 lysozyme by some 2.5–5.5 kJ mol^{-1} more than does the carboxyamide of Asn (Serrano *et al.* 1992*b*, *c*). Asp residues at positions N-cap + 1 and N-cap + 2, also stabilize the protein with respect to Asn residues by some 2.5–3.5 kJ mol^{-1} (Nicholson *et al.* 1991). Measurements on mutants that have histidine residues placed at both caps of both of the major helices of barnase shows the difficulties of distinguishing the effects of local hydrogen bonding from those of the macroscopic dipole but shows that the interaction energy between a single charge and the dipole is equal or greater than 2 kJ mol^{-1} (Sancho *et al.* 1992).

(e) *Amino acid residues at an internal position*

The changes in stability of barnase on making all 19 substitutions of Ala32 in the second helix (Horovitz *et al.* 1992) are in reasonable agreement with equivalent changes in internal positions in model helices (O'Neil & Degrado 1990; Lyu *et al.* 1990; Gans *et al.* 1991). Ala is the most stable residue while Gly destabilizes the helix relative to Ala by 2.8–3.2 kJ mol^{-1} and Pro by 12–16 kJ mol^{-1} (O'Neil & Degrado 1990; Horovitz *et al.* 1992; Sauer *et al.* 1992). There are several factors that could contribute to the differences in the intrinsic effect of different amino acids at internal positions on α-helix stability. First, burial of hydrophobic surface and van der Waals' contacts (Serrano *et al.* 1992*b*, *c*). Second, differences in the conformational restriction of the different amino acids in the helix relative to that in the unfolded proteins may account for most of the effects by contributing to the entropy of folding (Padmanabhan *et al.* 1990; Creamer & Rose 1992). Third, different solvation effects (Tobias & Brooks 1991). It must be emphasized that the surrounding residues and the position in the helix (Fairman *et al.* 1991; Pingchiang *et al.* 1991; Tobias & Brooks 1991; Serrano *et al.* 1992*b*, *c*) may provide the dominant effects on the contribution of a specific amino acid to helix stability.

6. Aromatic–aromatic interactions

Aromatic–aromatic interactions play an important role in the stereochemistry of organic reactions (Evans *et al.* 1987), binding affinities in host–guest chemistry (Jorgensen & Severance 1990) and probably also in protein stability and substrate binding affinities (Burley & Petsko 1985; Singh & Thornton 1985; Jorgensen & Severance 1990; Hunter & Sanders 1990; Hunter *et al.* 1991; Serrano *et al.* 1991). The contacts between phenylalanine rings in proteins occur with a significant higher

Downloaded from rsta.royalsocietypublishing.org

frequency than expected from a random distribution (Hunter *et al.* 1991). Two Tyr residues in a solvent exposed face of an α-helix in barnase (positions 13 and 17) are in an orientation that is close to the optimum for a tilted-T interaction. The magnitude of the direct interaction between two residues may be measured by the double-mutant cycle method (Serrano *et al.* 1990). The interactions of the two mutated groups with the rest of the protein tend to cancel out in the thermodynamic analysis and so the interaction between the two groups may be isolated. A double-mutant cycle analysis indicates that the interaction between Tyr13 and Tyr17 contributes about 5.2 kJ mol^{-1} to protein stability, and that the phenolic –OH group does not contribute to a significant extent to the aromatic pair interaction (Serrano *et al.* 1991). This value is about that expected for a hydrophobic interaction burying the same surface area. The aromatic rings of these residues make a multitude of interactions with other groups in the protein, and it is probably these that govern precise geometry of the interaction (Serrano *et al.* 1991).

7. Aromatic–histidine interactions

Protein engineering on barnase has uncovered the importance of an interaction between the imidazole ring of a histidine and the aromatic ring of Tyr, Phe or Trp where the protons from the histidine point towards the π-system of the aromatic ring (Loewenthal *et al.* 1992). Intriguingly, the interaction is stronger when the histidine is protonated. This contributes *ca.* 4 kJ mol^{-1} to protein stability at low pH. It also raises the pK_a of histidine by over half a unit, and this could be of general importance in modulating activity and binding (Loewenthal *et al.* 1992). Such interactions are discussed in detail by M. F. Perutz (this Volume).

8. Relation between statistical surveys and measured energetics

Professor J. Thornton has used empirical energetics derived from calculation from statistical data. If the probability of finding a particular residue, AA, at a particular position is P_{AA}, and if this represents true thermodynamic preference relative to say Ala (P_{Ala}), then a free energy may be calculated from the statistics($\Delta\Delta G_{stat}$):

$$\Delta\Delta G_{stat} = -RT \ln (P_{AA}/P_{Ala}). \tag{12}$$

The results of statistical survey may be compared with detailed energetic studies to see if this is valid. There is indeed a correlation between the statistical and observed experimental data ($\Delta\Delta G_{obs}$) for residues in α-helices with a correlation coefficient of 0.76 ($\Delta\Delta G_{obs} = 0.044 + 1.3221 \Delta\Delta G_{stat}$ (Serrano *et al.* 1992*a*)). These results indicate that the main factor in determining the statistical preferences is, indeed, thermodynamic.

9. Conclusions

All of the interactions within proteins are variable in magnitude and depend on context. This is because of tertiary and other extraneous interactions and local inhomogeneity in structure. The most important effect is that of hydrophobic interactions, especially within the core. In general, the magnitude of the hydrophobic effect is linearly related to the amount of accessible hydrophobic surface area that is buried to form the hydrophobic interaction. This can be severely modulated by specific packing interactions within a protein.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

References

Baker, E. N. & Hubbard, R. E. 1984 *Prog. Biophys. mol. Biol.* **44**, 97–179.

Burley, S. K. & Petsko, G. A. 1985 *Science, Wash.* **229**, 23–28.

Chothia, C. 1976 *J. mol. Biol.* **105**, 1–14.

Creamer, T. P. & Rose, G. D. 1992 *Proc. natn. Acad. Sci. U.S.A.* **89**, 5937–5941.

Daopin, S., Alber, T., Baase, W. A., Wozniak, J. A. & Matthews, B. W. 1991 *J. mol. Biol.* **221**, 647–667.

Dupré, A. 1869 *Théorie de la Chaleur*. Paris.

Eisenberg, D. & McLachlan, A. D. 1986 *Nature, Lond.* **319**, 199–203.

Eriksson, A. E., Baase, W. A., Wozniak, J. A. & Matthews, B. W. 1992 *Nature, Lond.* **355**, 371–373.

Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. 1992 *Science, Wash.* **255**, 178–183.

Evans, D. A., Chapman, K. T., Hung, D. T. & Kawaguchi, A. T. 1987 *Angew. Chem. int. Edn Engl.* **26**, 1184.

Fairman, R., Armstrong, K. M., Shoemaker, K. R., York, E. J., Stewart, J. M. & Baldwin, R. L. 1991 *J. mol. Biol.* **221**, 1395–1401.

Fersht, A. R. 1985 *Enzyme structure and mechanism*. New York: W. H. Freeman.

Fersht, A. R. 1987 *Trends Biochem. Sci.* **12**, 301–304.

Fersht, A. R. 1988 *Biochemistry* **27**, 1577–1580.

Fersht, A. R. & Serrano, L. 1993 *Curr. Op. struct. Biol.* **3**, 75–83.

Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y. & Winter, G. 1985 *Nature, Lond.* **314**, 235–238.

Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W. & Kallenbach, N. R. 1991 *Biopolymers* **31**, 1605–1614.

Hildebrand, J. H. 1979 *Proc. natn. Acad. Sci. U.S.A.* **76**, 194.

Hol, W. G. J. 1987 *Prog. biophys. mol. Biol.* **45**, 149–195.

Horovitz, A., Matthews, J. M. & Fersht, A. R. 1992 *J. mol. Biol.* **227**, 560–568.

Hunter, C. A. & Sanders, J. K. M. 1990 *J. Am. chem. Soc.* **112**, 5525–5534.

Hunter, C. A., Singh, J. & Thornton, J. M. 1991 *J. mol. Biol.* **218**, 837–846.

Hurley, J. H., Baase, W. A. & Matthews, B. W. 1992 *J. mol. Biol.* **224**, 1143–1159.

Jorgensen, W. & Severance, D. 1990 *J. Am. chem. Soc.* **112**, 4768–4774.

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

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

Lim, W. A., Farruggio, D. C. & Sauer, R. T. 1992 *Biochemistry* **31**, 4324–4333.

Loewenthal, R., Sancho, J. & Fersht, A. R. 1992 *J. mol. Biol.* **224**, 758–770.

Lyu, P. C., Liff, M. I., Marky, L. A. & Kallenbach, N. R. 1990 *Science, Wash.* **250**, 669–673.

Matsumura, M., Becktel, W. J. & Matthews, B. W. 1988 *Nature, Lond.* **334**, 406–410.

Matthews, B. W., Nicholson, H. & Becktel, W. J. 1987 *Proc. natn. Acad. Sci. U.S.A.* **84**, 6663–6667.

Nicholls, A., Sharp, K. A. & Honig, B. 1991 *Proteins Struct. Funct. Genetics* **11**, 281–296.

Nicholson, H., Anderson, D. E., Daopin, S. & Matthews, B. W. 1991 *Biochemistry* **30**, 9816–9828.

O'Neil, K. T. & Degrado, W. F. 1990 *Science, Wash.* **250**, 646–651.

Ooi, T., Oobatake, M., Nemethy, G. & Scheraga, H. A. 1987 *Proc. natn. Acad. Sci. U.S.A.* **84**, 3086–3090.

Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M. & Baldwin, R. L. 1990 *Nature, Lond.* **344**, 268–270.

Pingchiang, C. L., Wang, P. C., Liff, I. M. & Kallenbach, N. R. 1991 *J. Am. chem. Soc.* **113**, 3568–3572.

Preissner, R. & Bork, P. 1991 *Biochem. Biophys. Res. Commun.* **180**, 660–665.

Presnell, S. R., Cohen, B. I. & Cohen, F. E. 1992 *Biochemistry* **31**, 983–993.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Downloaded from rsta.royalsocietypublishing.org

Presta, L. G. & Rose, G. D. 1988 *Science, Wash.* **240**, 1632–1641.
Richards, F. M. & Richmond, T. 1978 *Ciba Symp.* **60**, 23–37.
Richardson, J. S. & Richardson, D. C. 1988 *Science, Wash.* **240**, 1648–1642.
Sancho, J., Serrano, L. & Fersht, A. R. 1992 *Biochemistry* **31**, 2253–2258.
Sandberg, W. S. & Terwilliger, T. C. 1991 *Proc. natn. Acad. Sci. U.S.A.* **88**, 1706–1710.
Sauer, U. H., Sun, D. P. & Matthews, B. W. 1992 *J. biol. Chem.* **267**, 2393–2399.
Serrano, L., Bycroft, M. & Fersht, A. R. 1991 *J. mol. Biol.* **218**, 465–475.
Serrano, L. & Fersht, A. R. 1989 *Nature, Lond.* **342**, 296–299.
Serrano, L., Horovitz, A., Avron, B., Bycroft, M. & Fersht, A. R. 1990 *Biochemistry* **29**, 9343–9352.
Serrano, L., Kellis, J. T., Cann, P., Matouschek, A. & Fersht, A. R. 1992*a* *J. mol. Biol.* **224**, 783–804.
Serrano, L., Neira, J. L., Sancho, J. & Fersht, A. R. 1992*b* *Nature, Lond.* **356**, 453–455.
Serrano, L., Sancho, J., Hirshberg, M. & Fersht, A. R. 1992*c* *J. mol. Biol.* **227**, 544–559.
Sharp, K. A., Nicholls, A., Fine, R. F. & Honig, B. 1991*a* *Science, Wash.* **252**, 106–108.
Sharp, K. A., Nicholls, A., Friedman, R. & Honig, B. 1991*b* *Biochemistry* **30**, 9686–9697.
Shirley, B. A., Stanssens, P., Hahn, U. & Pace, C. N. 1992 *Biochemistry* **31**, 725–732.
Shortle, D., Stites, W. E. & Meeker, A. K. 1990 *Biochemistry* **29**, 8033–8041.
Singh, J. & Thornton, J. M. 1985 *FEBS Lett.* **191**, 1–6.
Tobias, D. J. & Brooks, C. L. 1991 *Biochemistry* **30**, 6059–6070.
Van Gunsteren, W. F. & Mark, A. E. 1992 *Eur. J. Biochem.* **204**, 947–961.
Yutani, K., Ogasahara, K., Tsujita, T. & Sugino, Y. 1987 *Proc. natn. Acad. Sci. U.S.A.* **84**, 4441.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

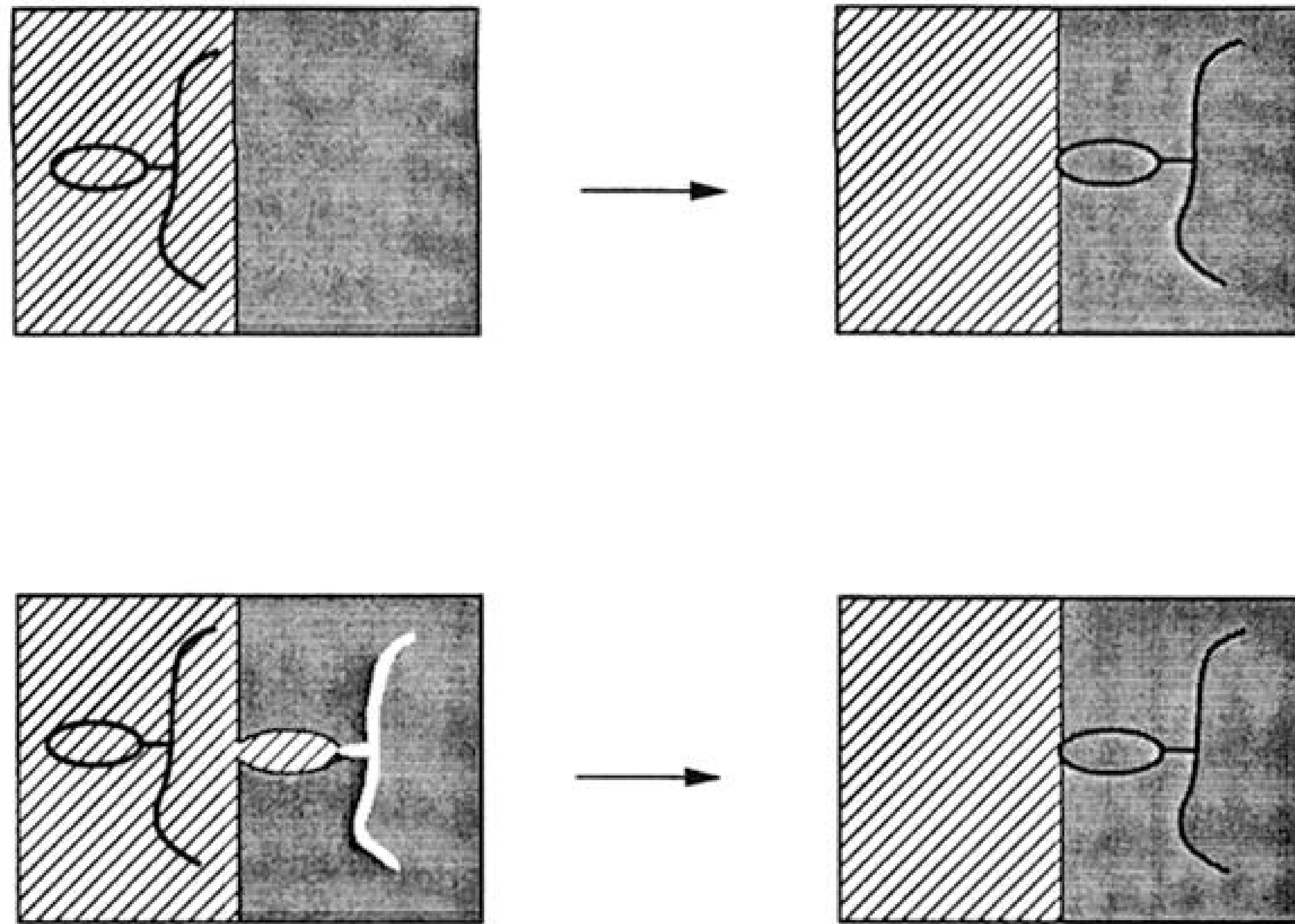

Figure 2. Cartoon comparing the transfer of a hydrophobic group side chain to a hydrocarbon and to a protein with a preformed cavity on the surface.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY