PREPRINT *Final version published in: "Protein: A Comprehensive Treatise"*
Volume 2, pp. 217-270 (1999)
Series Editor: Geoffrey Allen
Publisher: JAI Press Inc.

Thermodynamics of Protein Folding and Stability

Alan Cooper

Chemistry Department, Glasgow University
Glasgow G12 8QQ, Scotland, UK.

Phone: +44 (0)141-330 5278
FAX: +44 (0)141-330 2910
e-mail: alanc@chem.gla.ac.uk

In Memoriam: Christian B. Anfinsen (1916-1995) *

* Footnote: *ca.* 1971 I shared a rather dilapidated and now demolished office with Chris Anfinsen in South Parks Road, Oxford, during his sabbatical visit to the Molecular Biophysics Laboratory shortly before he won the Nobel Prize. Chris was a visiting fellow of All Souls College (or "Old Souls" as he usually liked to call it), and I was a still-wet-behind-the-ears postdoc. Memories of his charm, intellect, friendliness, and scientific humility have been a guiding influence ever since.

1. Introduction

Remarkable early work, notably by Hsien Wu and others (Wu, 1931; Anfinsen & Scheraga, 1975; Edsall, 1995), established the idea that denaturation of soluble proteins involved transitions from a relatively compact orderly structure to a more flexible, disorganized, open polypeptide chain. It was also known at this time that denaturation could be reversed. But it was the work of Anfinsen and colleagues in the late 1950s on the refolding of polypeptides that really galvanised interest in the physical chemistry of this process, particularly at the time when the molecular basis for the genetic code was being established (Anfinsen, 1973). The ability of polypeptides with appropriate primary sequence to fold into active native structures without, necessarily, the intervention of external agencies completes a vital link in the chain leading to expression of genetic information. Under the correct physicochemical conditions the folding of a protein is spontaneous and determined solely by its amino acid sequence. Once a gene is expressed, translated into a specific polypeptide sequence, thermodynamics (possibly guided by kinetics) takes over and the intrinsically flexible, irregular polymer chain folds into the more compact, specific structure required (usually) for biological function.

This ability for a polypeptide to select one conformation, spontaneously and usually quite rapidly, from a myriad of alternatives, has given rise to what has come to be called “The Protein Folding Problem”. This is really not just one problem but several, involving basic questions such as: How? Why? Whether? How a protein folds is a question (or series of questions) relating to mechanism. What are the pathways involved in the process whereby the unfolded protein (whatever that is) reaches the folded state ? What are the kinetics ? What intermediates are involved, if any, and are they unique ? What are the rate-limiting steps ? ...and so forth. It is an area which has become much more at the forefront recently with the demonstration of “chaperone” and related effects in protein folding. It is also of considerable interest to those attempting the awesome task of predicting protein structures from amino acid sequences, since the shortcuts taken by the protein itself may help in suggesting effective algorithms for predictive methods. However, these are treated more fully elsewhere in this series. Why a protein folds relates to the even more fundamental thermodynamic problem of the underlying molecular interactions responsible for stabilizing the folded conformation relative to other intrinsically more likely irregular states of the polypeptide. This is the subject to be covered here. Whether a protein folds depends on both the above. In order for a particular polypeptide sequence to adopt spontaneously a functionally effective conformation, the folded form must have a lower thermodynamic free energy than the galaxy of other available conformations. The folded conformation must also be kinetically attainable, with appropriate pathways, no unattainable intermediate states, and no irreversible kinetic traps.

My aim in this chapter is to review the thermodynamic background to protein folding and stability, with an overview of the current picture as I see it. Many detailed reviews in this area have appeared (Tanford,1968,1970;Privalov, 1979, 1982; Murphy & Freire, 1992), some of them very recently (Dill & Stigter, 1995; Honig & Yang, 1995; Lazaridis *et al*., 1995; Makhatadze & Privalov, 1995), and it is not my intention to cover the same ground in as much detail as can be found there. Rather, I will try to provide sufficient basic background to allow understanding and critical appraisal of this work by non-specialist readers.

1.1. Semantics: Definitions and General Considerations

Many of the conceptual difficulties in this field, especially for newcomers, arise from semantics: the way in which the same or apparently similar terminology is used to mean different things by different workers. Consciously or unconsciously, people with different backgrounds can use the same terms to mean entirely different things. And the definitions of terms may change over time as well, so the same terms encountered in some of the older literature may not carry the same meaning in more recent work. The term "random coil", for example, is a case in point. To a polymer chemist this might mean a highly flexible, dynamic, fluctuating, disordered chain structure in which no one molecule or region of a molecule is like any other. To a protein crystallographer however, this same term might be used to refer to those regions of a protein structure that do not contain any recognisable helix, sheet, or other motif - but yet is a quite fixed, well defined conformation identical from one molecule to the next. Because it is important not to be confused by conflicting terminology, in the next few sections I will try to clarify what I mean by the various possible conformational states of a polypeptide and the sorts of interactions that might be responsible for their occurrence.

1.1.1. Semantics I: Conformational States

Although polypeptides are inherently flexible polymers, we should be clear right from the very start that the "random coil" is the least likely state of any polypeptide in water. Free rotations about torsional angles (ϕ, ψ) of the peptide unit would allow a myriad of potential chain conformations[1]. But these rotations are by no means "free". Simple steric constraints, epitomized in the classic Ramachandran plot, restrict the range of realistically attainable ϕ-ψ angles even for a polypeptide in vacuum. The physical bulk of peptide atoms and sidechain groups prevents close encounters or overlap - except at a very high energy cost - and means that only relatively limited areas of ϕ-ψ space are available.

Moreover, polypeptide is intrinsically "sticky stuff" (one of the most abundant proteins, collagen, takes its name from the Greek κολλα = glue) and water is a far from ideal solvent. Hydrogen bonding of water molecules to peptide backbone -NH and -C=O groups will further restrict conformational freedom. Interactions, however transient, between peptide groups and side chain residues on the polypeptide will also take a part. (At higher concentrations, interactions between adjacent polypeptide molecules is also a factor of considerable importance, often leading to coagulation or aggregation of denatured proteins.)

Even so, the range of available conformations is enormous, and we must choose our language carefully when attempting to describe them.

Traditionally, emphasis is placed on the backbone conformations that a polypeptide might adopt, since these are easiest to describe. Hence if we could take a snapshot look at an individual polypeptide we might see differing amounts of:

Regular structure - involving a repeating pattern of ϕ-ψ angles, with defined H-bond connectivity, giving rise to the familiar α-helix, β-sheets, 3-10 helix.

---

[1] For a 100-residue protein, even allowing just 3 possible ϕ-ψ angles per peptide group would give rise to $3^{100} = 5 \times 10^{47}$ possible different conformations of the polypeptide chain. Such unimaginably large numbers gave rise to the "Levinthal paradox" (Levinthal, 1968; Dill, 1993) whereby there is insufficient time, even in the known lifetime of the universe, for any polypeptide to explore all these possibilities to find the "right" one.

Irregular structure - involving stretches of peptides with no repeating pattern of ϕ-ψ angles, and differing patterns of H-bonding, including hydrogen bonding to surrounding water molecules.

Motif structure - commonly occurring patterns of adjacent ϕ-ψ angles spanning just a few amino acids, not necessarily regular, but giving a recognisable conformational feature (e.g. β-bends, turns).

In a population of polypeptide molecules each of these structural classes might be:

Homogeneous - identical conformation in all molecules, with any one molecule superimposable upon another.

or

Heterogeneous - different conformations from one molecule to another, with different ϕ-ψ angles, H-bond connectivity, hydration, and so forth.

And this latter conformational heterogeneity might be:

Static - unchanging with time

or

Dynamic - changing randomly/stochastically with time in any one molecule.

[Similar considerations will apply equally to side chain conformations, though this is rarely done for reasons of complexity.]

It is worth emphasizing here that all protein molecules, whether folded or not, are dynamically heterogeneous - just like any other substance above absolute zero. On a short enough timescale, and over short enough distances:-

No part of any protein is ever static.
No protein molecule ever has exactly the same conformation as any other.
No protein molecule ever exists in the same conformation twice.

This is simply an unavoidable consequence of thermodynamics and the nature of heat (Cooper, 1976, 1984; Brooks *et al*. 1988), and might be pictured as just another manifestation of Brownian Motion at the (macro)molecular level. The timescale for dynamic fluctuations might be anything from femtoseconds to kiloseconds, and their experimental/functional consequences will depend on the relevant observational timescale. The magnitudes of the conformational fluctuations will be mostly small, involving thermal vibration, libration, torsion of individual groups, but much larger effects are also possible (Cooper, 1984).

Against this background, and given these definitions, how might we recognise or classify or define the different conformational states of a protein ? Maybe as follows:

Folded: - the biologically active ("native") form of the polypeptide (usually). Compact, showing extensive average conformational homogeneity with recognisable regions of regular, irregular and motif structures, on a background of dynamic thermal fluctuations. Well defined H-bond connectivity, much of it internalized, with secondary and tertiary structure characteristic of the particular protein.

Unfolded: - everything else ! An ill-defined state, or rather set of states comprising anything that is not recognisably folded. A population of conformations, spanning and sampling wide ranges of conformation space depending on conditions. Usually quite open, irregular, heterogeneous, flexible, dynamic structures - no one molecule is like another, nor like itself from one moment to another. But not necessarily "random coil" (see below) - some residual, transient secondary structure possible.

As sub-sets of the latter unfolded states we might have:-

Mis-folded: - Partially or incorrectly folded conformers, bearing some similarity to the native fold, but with regions of non-native, possibly heterogeneous structure. Might result from kinetic traps, or from chemical modification (proline isomerization, disulphide rearrangements, etc.).

Aggregated: - The classic "denatured", coagulated protein state. Intractable masses of entangled, unfolded polypeptide. The usual product of thermal unfolding of large proteins. Usually heterogeneous, but may contain regions of regular structure.

Molten Globule: - a relatively compact, globular set of conformations with much regular, secondary structure in the polypeptide backbone, but side chain disorder. First characterized by affinity for hydrophobic probes - popular candidates as intermediates in the folding pathway (Ptitsyn, 1995; Privalov, 1996). [Caution: not all workers agree on a definition for "molten globule"!]

Random Coil: - this is the (hypothetical) state in which the conformation of any one peptide group is totally uncorrelated with any other in the chain, particularly its neighbours. All polypeptide conformations are equally likely, equally accessible, and of equal energy. Populations of such molecules would show complete conformational heterogeneity. This state is almost certainly never found for any polypeptide in water ! (Though, unfortunately, the term is sometimes usurped by protein crystallographers to describe the regions of their structures - loops, etc. - that are not immediately identifiable as any of the regular structures or motifs. These are best described as irregular structure - and may be homogeneous or heterogeneous, static or dynamic, depending on circumstances.)

1.1.2. Semantics II: Interactions

Another semantic minefield is encountered when considering the forces responsible for biomolecular interactions. Although in principle the energy of any state of a macromolecular system should be obtainable by solution of the appropriate quantum mechanical (Schrödinger) equations, in practice such an approach is not yet practicable except in very special and well-defined circumstances. And, even if feasible, such calculations would be conceptually unhelpful and would lack the thermodynamic dimension that might relate derived parameters to experimental observables. In such a situation it has been traditional to be guided by analogy and experience from other areas of physical chemistry of (generally) small molecules, and attempt to break down the overall interaction into discrete categories of pair-wise interactions between recognisable molecular groupings. This is the origin of more-or-less familiar terms such as: “bonded”, “non-bonded”, “non-covalent”, “polar”, “electrostatic”, “hydrogen bond”, “hydrophobic”, “solvation”, “van der Waals”, “dispersion” - and more - interactions.

Bonded interactions are usually considered to be those directly involved in the covalent links between adjacent atoms. Stretching, bending, or rotation of these bonds, either in the polypeptide backbone or sidechain groups, will require work and will change the total energy of the system. Covalent bond stretching or bending is particularly hard work and requires energies that are usually beyond the normal range for thermal motions. Consequently it is usually assumed that covalent bonds in proteins adopt their minimum energy, least strained conformations (bond lengths and angles) wherever possible. Except for the peptide group, however, rotation about many covalent bonds is relatively easy, and this is the source of inherent flexibility in the unfolded polypeptide.

Non-bonded or non-covalent interactions are those between atoms or groups that are separated by more than one covalent bond. Confusingly, such interactions may be referred to as being “short-range” or “long-range”, either in terms of the through-space distances between groups or, frequently, in terms of separation in sequence along the polypeptide chain. Consequently, a non-covalent interaction between two amino acid residues might be “long-range” if the residues are separated by long stretches of polypeptide in the primary sequence, yet at the same time “short-range” if, through folding, the groups lie next to each other in space.

Non-covalent interactions may be broken down into the familiar categories listed above. Although it is not possible to give more than a qualitative description of the thermodynamic characteristics of each of these interaction categories at this stage, a brief description here might be useful. More details will emerge later in discussion of the folding problem.

Van der Waals or London dispersion forces are the ubiquitous attractive interactions between all atoms and molecules that arise from quantum mechanical fluctuations in the electronic distribution. They are consequences of the Heisenberg uncertainty principle. Transient fluctuations in electron density distribution in one group will produce changes in the surrounding electrostatic field that will affect adjacent groups. In the simplest picture, a transient electric dipole will polarise or induce a similar but opposite dipole in an adjacent group such that the two transient dipoles attract. The dipole-dipole interaction is truly short range, varying as inverse 6th. power of the separation distance, and such interactions are usually only of significance for groups in close contact. The strength of the interaction also depends on properties such as high-frequency polarizability of the groups involved, but apart from this, such interactions involve very little specificity. All atoms or groups will show van

der Waals attractions for each other. Also sometimes included in van der Waals interactions is the very steep repulsive potential between atoms in close contact ("van der Waals contact"). This arises from repulsions between overlapping electronic orbitals in atoms in non-covalent contact which makes atoms behave almost like hard, impenetrable spheres at sufficiently short range. Thermodynamically, van der Waals interactions would normally be considered to contribute to the enthalpy of interactions, with no significant entropy component.

Permanent dipoles and charges within molecules or groups give rise to somewhat longer range and more specific electrostatic interactions. Discrete charge-charge or dipole-dipole interactions may be attractive or repulsive, depending on sign and orientation. A particularly close, direct electrostatic interaction between ionized residues in a structure might be called a "salt bridge". Permanent dipoles or other electronic distributions may also polarise surrounding groups to give static induced dipoles, etc., that may interact attractively. The complete description of the electrostatics of the polypeptide, folded or otherwise, must also take into account interactions with surrounding solvent water molecules and other ionic species in solution. This means that thermodynamic description of such interactions is complicated and includes both enthalpy and entropy terms. For example, even the apparently simple process of dissolving of a crystalline salt in water can be endothermic or exothermic, depending on ion size and other factors, and can be dominated by entropic contributions from solvation, restructuring of water around ions, or other indirect effects not normally visualised in the simple pulling apart of charged species. Comprehensive studies of protein and related electrostatics are described by Honig *et al*. (1993).

Hydrogen bonds are now normally considered to be examples of particularly effective electrostatic interaction between permanent electric dipoles, especially in proteins between groups such as -NH and -C=O or -OH, and the -NH---O=C- interaction is of particular historical importance for the part it played in predictions of regular helical or sheet conformations. In theoretical calculations H-bond interactions may be handled either discretely as separate "bonds" or incorporated into the overall electrostatics of the protein. The thermodynamic contribution of hydrogen bonds to protein stability or other biomolecular interactions is surprisingly unclear. And the term "strength of the hydrogen bond" is very ambiguous. This is because liquid water is a very good hydrogen bonding solvent. Breaking of a hydrogen bond between two groups in a vacuum requires a significant amount of energy - in the region of 25 kJ mol^{-1} for a peptide hydrogen bond, say (Rose & Wolfenden, 1993; Lazaridis *et al*., 1995). But in water, such exposed groups would likely form new H-bonds to surrounding water molecules to cancel the effect, and the true "strength of a hydrogen bond" between groups in an aqueous environment might be closer to zero. The overall interaction will also include significant entropy contributions because of this solvent involvement. The usually excellent solubility of polar compounds in water reflects this, and model compound studies generally lead to a picture in which hydrogen bonds contribute little if anything to the free energy of folding of a polypeptide chain (Klotz & Farnham, 1968; Kresheck & Klotz, 1969; and others, see Dill, 1990). [They will, of course, determine the specific conformation adopted by the polypeptide when it does fold.]

Hydrophobic interactions are another manifestation of the peculiar hydrogen bonding properties of water. Based on empirical observation that non-polar molecules are poorly soluble in water, this interaction is probably best visualised as a repulsive interaction between non-polar groups and water, rather than a direct attraction between those groups. Non-polar, hydrophobic groups in water will tend to cluster together because of their mutual repulsion from water, not necessarily because the have any particular direct affinity for each

other. The thermodynamics of this interaction are interesting (Kauzmann, 1959; Tanford, 1980). Based on studies of small non-polar molecules, the separation or pulling apart of two hydrophobic groups in water is an exothermic process. In other words, although it generally requires work to separate such groups, heat is given out in the process. This is usually described in terms of structural rearrangements of water molecules at the molecular interface - but the molecular description is really less relevant than the empirical observation. This exothermic effect is opposed by a significant and thermodynamically unfavourable reduction in entropy of the system, also attributable to solvent structure rearrangements. The reverse process, that is the association of non-polar groups to form a "hydrophobic bond" , is consequently said to be "entropy driven" and comes about spontaneously even though it is endothermic. The enthalpies or heats of such processes are also characteristically temperature dependent (ΔC_p effect - see later), and this has been some of the stronger evidence for the role of such interactions in protein folding.

1.2. Thermodynamics

We know from experience that transformation of a protein between various conformational states might be brought about by changes in temperature, pressure, pH, ligand concentration, chemical denaturants or other solvent changes. For each of these empirical variables there will be a set of associated thermodynamic parameters, and it is axiomatic (Le Chatelier's Principle) that a transformation may only come about if the two states have different values for these parameters. For example, temperature-induced protein unfolding (at equilibrium) arises from differences in enthalpy (ΔH) between folded and unfolded states; pressure denaturation can only occur if the folded and unfolded states have different partial molar volumes (the unfolded state is normally of lower volume); unfolding at high or low pH implies differences in pK_A of protein acidic or basic groups; ligand-induced unfolding or stabilization of the native fold results from differences in binding affinity for ligand in the two states; chemical denaturants may act as ligands, binding differently to folded or unfolded states, or may act indirectly via changes in overall solvent properties. In each of these cases we need to know how to measure and interpret these thermodynamic parameters.

One important observation is that the "folded <--> unfolded" transition is highly co-operative, at least for small globular proteins, frequently behaving as an almost perfect 2-state equilibrium process akin to a macroscopic phase change (see Dill 1995). This feature will be discussed in some more detail later. But our task here is to describe how the thermodynamics of transition between these various states may be measured and interpreted, leading to a possible understanding of why the native folded form is usually the more stable state under relevant conditions. The arguments must necessarily be thermodynamic. We have already had cause to use terms such as "enthalpy", "entropy", "free energy" - and it is important to be clear what these terms mean. Experts in thermodynamics may skip the next section.

1.2.1. Basic Thermodynamics: A Primer

Thermodynamics can be a daunting subject. For that reason it is perhaps useful to summarise here the basic concepts, presented in a somewhat less conventional manner than found in the usual textbooks. What follows is a very unrigorous and highly abbreviated sketch of basic ideas of "molecular thermodynamics" or "statistical mechanics", starting from a molecular point of view and leading to classical thermodynamic relations. My aim is to encourage basic

understanding of thermodynamic expressions in a way that may make standard texts more readable to the non-expert.

Except at absolute zero, all atoms and molecules are in perpetual, chaotic motion. Things we feel, like "heat" and "temperature", are just macroscopic manifestations of this motion. Although in principle one might think it possible to calculate this motion exactly (using Newton's laws of motion or quantum mechanical equivalents), in practice this is impracticable for systems containing more than just a few molecules over a realistic timescale, and downright impossible for macroscopic objects containing of order 10^{23} molecules. And in any case, the information given by such a calculation would be far too detailed to be of any real use.

The way out of this problem is to take a statistical approach (statistical mechanics or thermodynamics) and concentrate on the average or most probable behaviour of the molecules. This will give the mean properties, what we observe for a sample containing large numbers of molecules, or the time-averaged behaviour of a single molecule.

The basic rule - a paraphrase of the Second Law of Thermodynamics at the molecular level - is that: The Most Probable Things Generally Happen.

The statistical probability (p_A) that any molecule or system (collection of molecules) is to be found in some state, A, depends on the energy (E_A) of the system together with the number of ways (w_A) that energy may come about. This is expressed in the Boltzmann probability formula:

$$p_A = w_A.\exp(-E_A/k_BT)$$

where T is the absolute temperature (in Kelvin), k_B is Boltzmann's constant ($k_B = 1.38 \times 10^{-23}$ J K^{-1}) and, again, E_A is the total energy of the system, comprising all the molecular kinetic, rotational and vibrational energy, together with energy due to interactions ("bonds") within and between the molecules in the system, and w_A is the number of ways in which that total energy may be achieved or distributed.

Some points of detail now need to be taken into account. Firstly, it is conventional and convenient to think in terms of <u>moles</u> of molecules rather than actual numbers of molecules in the system. Therefore we may multiply numerator and denominator of the energy exponent ($-E_A/k_BT$) by Avogadro's number (N_A), remembering that the gas constant $R = N_Ak_B = 8.314$ J K^{-1} mol^{-1} and redefining E_A as the total energy <u>per mole</u>, to give $-E_A/RT$ in the exponential factor. Secondly, since most of the time we work under conditions of constant pressure, we need to make sure that the energy accounting is properly formulated to take account of any work terms arising from volume changes (to satisfy energy conservation, or the First Law of Thermodynamics). This is done by taking <u>enthalpy</u> (H_A) as the appropriate energy term. Formally the enthalpy of a system is defined:

$$H_A = U_A + PV$$

where U_A is the internal energy, comprising molecular kinetic, rotational, vibrational and interaction energies in the system, and the pressure-volume term (PV) takes care of any energy changes due to work done on or by the surroundings.

Putting these points together leads to an equivalent version of the Boltzmann probability factor:

$$p_A = w_A.\exp(-H_A/RT)$$

Now consider a situation where our system might also exist in another state B, say, with probability

$$p_B = w_B.\exp(-H_B/RT)$$

and is free to interconvert between the two. We might depict this chemically as:

$$A \rightleftharpoons B$$

For a large number population of molecules in the system (or for smaller numbers averaged over a period of time) the relative probability of finding the system in either state is equal to the conventional "equilibrium constant" (K) for the process:

$$K = [B]/[A] = p_B/p_A \qquad \text{(where [] implies molar concentration)}$$

consequently, using the Boltzmann probability terms and after a little rearrangement we might write:

$$-RT.\ln(K) = \Delta H^O - RT.\ln(w_B/w_A)$$

where $\Delta H^O = H_B - H_A$ is the (molar) enthalpy difference between the two states.

This is equivalent to the classical thermodynamic expressions[2]:

$$\Delta G^O = -RT.\ln(K) = \Delta H^O - T.\Delta S^O$$

provided we identify $\Delta S^O = R.\ln(w_B/w_A)$.

In other words:

(i) The "standard Gibbs Free Energy change" (ΔG^O) is just another way of expressing the relative probability (p_B/p_A) of finding the system in either state. If ΔG^O is positive, $p_B/p_A <$

[2] For technical reasons, the superscript zeros in ΔG^O and ΔS^O are important - they designate changes occurring under standard state conditions. In the simple $A \rightleftharpoons B$ isomerization example here only the concentration ratios matter, not their absolute values. But in more general cases, where the number of molecules can change during reaction, we must correct for entropy of mixing contributions or relate everything to defined standard states. In contrast, the variation in enthalpy with concentration is normally insignificant, and it is usually permissible to use ΔH and ΔH^O interchangeably. See any standard thermodynamics text for details.

1, and state B is relatively unlikely. If ΔG^O is negative, $p_B/p_A > 1$, and state B is the more likely. When $\Delta G^O = 0$, $p_A = p_B$, and either state is equally likely (or the equivalent, the system spends 50% of its time in either state).

(ii) The "standard entropy change" (ΔS^O) is just an expression of the change in the different numbers of ways in which the energy of the system in a particular state may be made up.

It is this latter which helps (me, at least) get a better feeling for the concept of entropy. Following Boltzmann, the absolute molar entropy of any system is given by: $S = R.\ln(w)$, and is just a way of expressing the multiplicity of ways in which the system can be found with a particular energy, sometimes called the "degeneracy" of the system. [Elementary descriptions of entropy couched in terms of "randomness" or "disorder" can be confusing or ambiguous - for example, the distribution of symbols on this page might look somewhat random to someone who cannot read, but there is really only one way (or relatively few ways) that make sense.]

It is important to emphasise that the most probable (equilibrium) state of a system is determined by the Gibbs Free Energy, reflecting the relative probabilities, and that this is made up of a combination of energy (enthalpy) and entropy terms. Consequently, spontaneous processes need not necessarily involve a decrease in internal energy/enthalpy. Endothermic processes are quite feasible, indeed common (e.g. the melting of an ice cube at room temperature) provided they involve a suitably large increase in entropy.

The exponential nature of the Boltzmann probability expression seems to imply that low energy states are more likely and that things should tend to roll downhill to their lowest energy (enthalpy) state, as they do in conventional mechanical systems. And, all things being equal, that is what happens thermodynamically too. However, this is generally offset by the "w" term. The higher the energy, the more ways there are of distributing this energy in different ways to reach the same total. Except in special cases, the very lowest energy state of any system has all molecules totally at rest in precise locations (on lattice sites, for example) and there is generally only one way that this can be done ($w = 1$, $S = 0$). For higher energy states, however, there will be more ways in which that energy can come about - some molecules might be rotating, others vibrating, others moving around in different directions, some forming hydrogen bonds, others not, and any combination of these in multiple ways to make up the same total energy - indeed the way in which the total energy is distributed will vary with time as a result of molecular collisions, and the higher the energy the greater the number of ways there might be of achieving it. Expressed graphically (Fig.1), the decreasing exponential energy term combined with the increasing w component means that the most probable, average energy of any system is not the ground state (except for $T = 0$ K).

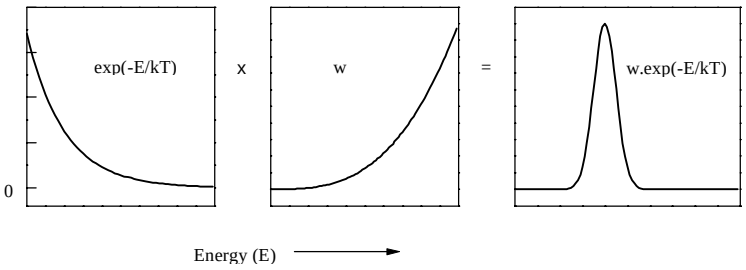

Fig.1: Graphical illustration of how the combination of exponentially decreasing Boltzmann factor, combined with rapidly increasing degeneracy (w), gives an energy probability distribution of finite width peaking at energies above zero.

1.2.2. Heat capacity

Both enthalpy and entropy are classical concepts related to the heat uptake or heat capacity of a system. Imagine starting with an object at absolute zero (0 K) in its lowest energy state. As we add heat energy, the temperature will rise and the molecules will begin to move around, bonds will break, and so forth. The amount of heat energy required to bring about a particular temperature increment depends on the properties of the system, but is expressed in terms of the heat capacity. At constant pressure, the heat energy (dH) required to produce a temperature increment dT is given by

$$dH = C_p.dT$$

where C_p is the heat capacity of the system at constant pressure. [Similar expressions are available for constant volume situations, but these are rarely encountered in biophysical experiments.]

Consequently, the total enthalpy of a system in a particular state at a particular temperature is simply the integrated sum of the heat energy required to reach that state from 0 K:

$$H = \int_0^T C_p.dT + H_0$$

where H_0 is the ground state energy (at 0 K) due to chemical bonding and other non-thermal effects.

The magnitude of the heat capacity (C_p) depends on the numbers of ways there are of distributing any added heat energy to the system, therefore is related to entropy. Consider the energy required to bring about a 1 K rise in temperature, say. If a particular system has only relatively few ways of distributing the added energy (w small, entropy low), then relatively little energy will be required to raise the temperature, and such a system would have relatively low C_p. If, however, there are lots of different ways in which the added energy can be spread around amongst the molecules in the system (w high, entropy high), then much more energy will be needed to bring about the same temperature increment. Such a system would have a high C_p.

This is expressed in the classical (2nd. Law) definition of an entropy increment (at constant pressure):

$$dS = dH/T = (C_p/T).dT$$

so that the total entropy of any system is given by the integrated heat capacity expression:

$$S = \int_0^T (C_p/T).dT$$

It is these equations, and variants below, connecting both enthalpy and entropy to heat capacity measurements, that make calorimetric methods potentially so powerful in determining these quantities experimentally in an absolute, model-free manner - see later.

When defined in this way, these quantities are absolute enthalpies and entropies of the system relative to absolute zero. But we are normally interested in changes in these quantities (ΔH, ΔS) from one state to the other at constant temperature (or over a limited range of temperatures close to physiological). These follow directly from the integral expressions above:

$$\Delta H = H_B - H_A = \int_0^T \Delta C_p .dT + \Delta H(0)$$

$$\Delta S = S_B - S_A = \int_0^T (\Delta C_p /T).dT$$

where $\Delta C_p = C_{p,B} - C_{p,A}$ is the heat capacity difference between states A and B at a given temperature. $\Delta H(0)$ is the ground state (0 K) enthalpy difference between A and B. Most systems are assumed to have the same (zero) entropy at absolute zero (3rd. Law of Thermodynamics).

It is frequently convenient to relate these quantities to some standard reference temperature T_{ref} (e.g. T_{ref} = 298 K rather than 0 K), in which case:

$$\Delta H(T) = \Delta H(T_{ref}) + \int_{Tref}^T \Delta C_p .dT$$

and $$\Delta S(T) = \Delta S(T_{ref}) + \int_{Tref}^T (\Delta C_p /T).dT$$

This emphasises that, if there is a finite ΔC_p between two states, then ΔH and ΔS are both temperature dependent - this is the norm when weak, non-covalent interactions are involved, and is particularly true for protein folding transitions. [This effect is generally less significant - at least over limited temperature range - for conventional chemical reactions, involving

covalent bond changes where large energy difference between the two chemical states are manifest even at absolute zero by differences in ground state energy.]

If ΔC_p is constant, independent of temperature (not necessarily true, but usually a reasonable approximation over a limited temperature range), then we can integrate the above to give approximate expressions for the temperature dependence of ΔH and ΔS with respect to some arbitrary reference temperature (T_{ref}):

$$\Delta H(T) = \Delta H(T_{ref}) + \Delta C_p .(T - T_{ref})$$

$$\Delta S(T) = \Delta S(T_{ref}) + \Delta C_p .\ln(T/T_{ref})$$

showing how ΔH and ΔS will both vary with temperature in the same direction. Thus, if ΔC_p is positive, both ΔH and ΔS will together increase with temperature in line with intuition - a higher enthalpy implies higher molecular energy states, broken bonds, and the like, consistent with higher entropy, greater degeneracy of the system. Similarly, lower entropy states are usually associated with more ordered systems with concomitantly lower enthalpy.

These synchronous changes in ΔH and ΔS with temperature tend to complement and cancel each other in the ΔG term, so the resulting changes in ΔG are significantly less. For example, for small changes in temperature $\delta T = T - T_{ref}$, using the approximation $\ln(1 + x) = x$, for $x \ll 1$:

$$\Delta H(T) = \Delta H(T_{ref}) + dCp.\delta T$$

$$\Delta S(T) = \Delta S(T_{ref}) + \Delta C_p .\ln(1 + \delta T/T_{ref}) \approx \Delta S(T_{ref}) + \Delta C_p .\delta T/T_{ref}$$

so the ΔC_p terms will partly (though not completely) cancel in ΔG.

Moreover, over a limited temperature range for which this approximation is valid:

$$\Delta H(T) \approx \Delta H(T_{ref}) + T_{ref}.(\Delta S - \Delta S(T_{ref}))$$

so that a plot of ΔH versus ΔS would appear linear with slope T_{ref}. Though much could be made of the significance of such a linear correlation, and the nature of T_{ref} as some sort of "characteristic temperature", it is simply a mathematical consequence arising from experimental data covering a limited temperature range. The T_{ref} arising from such a correlation would simply be that temperature for which the approximation (δT small) is most appropriate, i.e. somewhere in the experimental observable range.

These effects are one example of the much broader phenomenon of "enthalpy-entropy compensation" (Lumry & Rajender, 1970; Grunwald & Steel, 1995; Dunitz, 1995 - and references therein) whereby ΔH and ΔS changes brought about by various experimental conditions (in addition to temperature) tend to move in concert in such a way as to cancel almost quantitatively in ΔG. Much has been made of this in terms of special solvent/water properties, and so forth, but it is almost certainly just a simple manifestation of the intuitively

reasonable properties of systems comprising multiple, weak, non-covalent interactions as described above - high enthalpy implies high entropy, and *vice versa* (Weber, 1993, 1995; Dunitz, 1995).

1.2.3. The van't Hoff Enthalpy/Equation

The temperature dependence of the equilibrium constant for any process is a manifestation of the enthalpy of the process and forms the basis for widely used methods for estimating ΔH. Given that:

$$-RT.\ln K = \Delta H^{O} - T.\Delta S^{O}$$

then $$\ln K = -\Delta H^{O}/RT + \Delta S^{O}/R$$

and $$d(\ln K)/d(1/T) = -\Delta H^{O}/R$$

{Note: this is true whether or not ΔH^{O} and ΔS^{O} vary with temperature. In general:

$$d(\ln K)/d(1/T) = -\Delta H^{O}/R - (1/RT)[d(\Delta H^{O})/(1/T)] + (1/R)[d(\Delta S^{O})/d(1/T)]$$
$$= -\Delta H^{O}/R$$

since: $$d(\Delta H^{O})/d(1/T) = -T^{2}.d(\Delta H^{O})/dT = -T^{2}.\Delta C_p$$
and: $$d(\Delta S^{O})/d(1/T) = -T^{2}.d(\Delta S^{O})/dT = -T^{2}.\Delta C_p/T$$

so the latter two terms cancel in the above equation.}

As a consequence, a plot of experimental data of lnK vs. 1/T ("van't Hoff plot") gives a line whose slope at any point is the van't Hoff enthalpy (ΔH^{O} or ΔH_{VH}) divided by R. In simple cases, over a limited temperature range, this plot is linear (or is assumed to be so), but in general the temperature dependence of ΔH (due to ΔC_p) will result in a curved van't Hoff plot that needs more careful analysis (Naghibi *et al*. 1995). In practical terms the analysis can be made even more complicated (and such methods less satisfactory for ΔH determination) by the natural tendency described above for ΔH and ΔS to vary with temperature in a complementary manner so as to cancel and give relatively smaller changes in ΔG.

What is a "van't Hoff Enthalpy" ? To what does this energy refer ? It is important to recognise that any van't Hoff analysis is based on a model or assumption about the process involved. Typically this will be a "2-state" model (see below) in which the equilibrium constant K is a dimensionless ratio determined, usually indirectly, from spectroscopic, calorimetric, or other measurements. In such a model the molar van't Hoff enthalpy change, ΔH_{VH} , is the enthalpy change per mole of cooperative unit (Sturtevant, 1974). More on this later.

1.3. Thermal Energies and Fluctuations

Since all molecules are always in perpetual thermal motion (and thermodynamics is merely a consequence of this) it is useful to bear in mind the average thermal energies involved in

such motion. Classical statistical thermodynamics (the "equipartition theorem") show that every independent form of motion, or degree of freedom in a molecule has a mean thermal energy of $\frac{1}{2}k_BT$, where k_B is Boltzmann's constant and T is the absolute temperature. For kinetic energy or translational motions there are three degrees of freedom, corresponding to movement along *xyz* axes, so average kinetic energy is $3k_BT/2$. Similarly for free rotational motion the average energy will be about $\frac{1}{2}k_BT$ per rotational degree of freedom. Vibrational modes have two degrees of freedom each - one translational and one extensional - but for covalent bonds at least the classical equipartition approximation breaks down. Quantization of vibrational levels has to be considered here and conventional bond vibrations are rarely excited at normal temperatures. However, soft modes with frequencies of order $300cm^{-1}$ or less, such as might be found in global protein vibrations, will be thermally populated at physiological temperatures.

A useful rule of thumb is that the average thermal energy associated with each motional degree of freedom in a molecule is of order k_BT per molecule, or RT per mole. This corresponds to about 2.5 kJ mol^{-1} (0.6 kcal mol^{-1}) at room temperature.

There is another consequence of the statistical description of thermodynamics apparent from Fig.1. As with any statistical distribution, the energy probability of any system will have a finite width, and we should expect to see statistical fluctuations about the mean or most probable value. For large systems the distribution is usually very sharp, and fluctuations are not normally perceptible. But as systems get smaller, thermodynamic fluctuations get comparatively larger, as in Brownian motion, for example. For very small systems such as an individual protein molecules, the thermodynamic energy and volume fluctuations can be significant and play a definite role in the dynamic functions of the protein (Cooper 1976, 1984).

1.4. The 2-State Approximation

Many experimental methods for estimating thermodynamic parameters for protein transitions rely on the assumption/approximation of "2-state" behaviour for the system. The accuracy of the data thus obtained, and the validity of their interpretation are critically dependent on the validity of this assumption.

The 2-state model assumes that the process of interest (or part of it) may be represented by a thermodynamic equilibrium between two experimentally distinguishable states:

$$A \rightleftharpoons B$$

with no significant population of intermediate states and/or, equivalently, a relatively high kinetic activation barrier between them.

This does not necessarily imply that A and B themselves are unique, homogeneous, static states. Consider an ice cube at $0\ ^{O}C$, for example. This is a classic example of a macroscopic phase transition described extremely well by the 2-state approximation. The system can exist in one of two macroscopically distinguishable states: solid (ice) or liquid (water). At $0\ ^{O}C$ and 1atm pressure these two states can coexist, and the equilibrium can be shifted one way or

the other by slight changes in temperature, pressure, or composition of the system (additives). There are no known intermediates - nothing halfway between solid and liquid. At the molecular level the ice --> water melting transition, brought about say by increase in temperature, is characterized by a breaking of (some) intermolecular hydrogen bonds and loss of regular crystal lattice structure. However, not all H-bonds are broken. Estimates differ, and it is not even clear that the term “broken hydrogen bond” is useful for the description of interactions of water molecules in the pure liquid (Eisenberg & Kauzmann, 1969), but of order 50% remain unbroken at 0 ^{O}C. Further increase in temperature involves progressively further breaking of water-water H-bonds in the liquid (until eventually they all break and we have a second 2-state transition: boiling). Consequently, state B (liquid water) in this case is not a unique state but a continuum of states that merge smoothly and non-cooperatively with, if we could see them, differing average structures, extent of H-bonding and other properties. Similarly, the solid ice phase (state A) will vary with temperature - progressive changes in numbers of lattice defects, thermal disorder, vibrational amplitudes, lattice spacing (due to thermal expansion thermal expansion), and so forth. For example, root-mean square amplitudes of thermal vibration of atoms in ice I increase from 0.09 to 0.215 Å (for O atoms) or from 0.15 to 0.25 Å (for H atoms) over the -273 to 0 ^{O}C temperature range (Eisenberg & Kauzmann, 1969, p.78).

In the case of proteins, A and B might be the “native” (N) and “unfolded” (U) states, respectively, and the transition may be brought about by changes in temperature, pH or denaturant concentration. The U state does not, necessarily, have to become random coil, nor even fully unfolded during the 2-state transition, and might continue to change - become “more unfolded” - as more denaturant is added, or higher temperatures reached, for example.

The important experimental criterion is that there be some perceptible change in some observable property of the system that we might take as measure of the extent of the transition. For our lump of ice this might be volume, fluidity, calorimetric enthalpy, etc. For a protein this might be fluorescence, UV absorbance (reflecting environmental changes of aromatic groups), circular dichroism (CD), NMR parameters, calorimetric enthalpy, or others. In any case, experimentally we would measure some quantity (F) whilst varying some parameter (x), which might be temperature, pressure, denaturant concentration, etc., and expect to see sigmoidal variation typical of a 2-state transition (Fig.2).

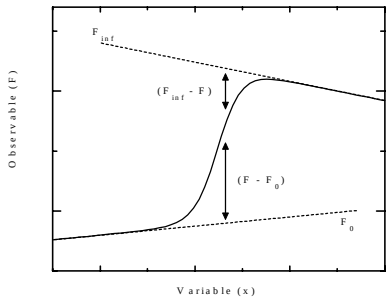

Fig.2: Illustration of sigmoidal variation of an experimental observable (F) with changing parameter (x) for a two-state transition, including pre- and post-transition baseline slopes.

The pre- and post-transition baseline slopes reflect the earlier argument that the properties of A and B themselves are expected to vary with x. After suitable correction for this, usually by linear extrapolation, the 2-state assumption allows us to estimate the (apparent) equilibrium constant (K_{app}) as a function of x:

$$K_{app}(x) = [B]/[A] = (F - F_0)/(F_{inf} - F)$$

where the square brackets [] indicate molar concentrations (strictly activities), F is the observed quantity, and F_0 and F_{inf} are the (extrapolated) values at low and high x values, representing pure A (N) or pure B (U) states respectively.

If the experimental variable is the temperature (T), then such data, giving K_{app} as a function of T, may be used to estimate the van't Hoff enthalpy change (ΔH_{VH}) for the transition.

It is pertinent to consider again what is meant by the "van't Hoff enthalpy" in these circumstances and, in particular, how it depends on the size of the system undergoing the 2-state transition.

Note that K_{app} is a dimensionless quantity, and that we do not normally need to know the absolute concentrations of A and B in order to determine it - simply the ratio of appropriate F values is sufficient. Yet ΔH_{VH} has the units kJ per mole (or equivalent, i.e. the units of R in the van't Hoff equation). Per mole of what, we may ask ? Well, it is per mole of whatever is undergoing the 2-state transition, or per mole of the "cooperative unit". This depends on the size of the system. For our block of ice this would be the enthalpy change for a mole of (identical) ice cubes - since it is the whole ice cube that melts cooperatively. For a protein molecule (or, more strictly, a solution of protein molecules) we might anticipate the cooperative unit to be just the molecule itself since, although individual molecules might unfold cooperatively, the behaviour of separate molecules is uncorrelated.

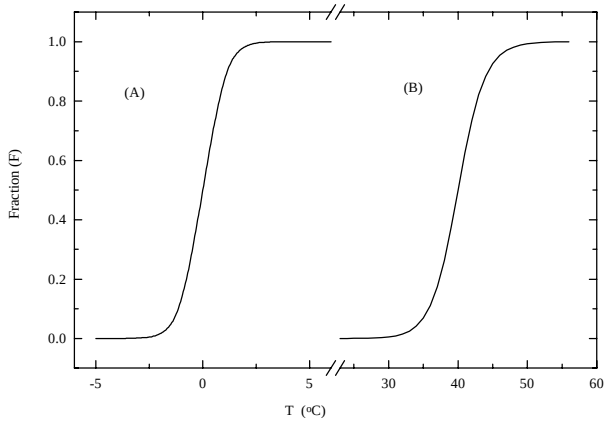

Fig.3: Sigmoidal van't Hoff transition curves showing fractional extent of the transition (F) versus temperature (T) for: (A) a hypothetical ice cube, 20 Å per side; (B) a typical protein molecule unfolding at 40 ^{O}C with $\Delta H_{VH} \approx$ 400 kJ mol^{-1} (ca. 100 kcal mol^{-1}).

This is illustrated in Fig.3 showing the (sigmoidal) transition with increasing temperature expected for 2-state van't Hoff behaviour for ice compared to a typical protein with $\Delta H_{VH} \approx$ 400 kJ mol^{-1} .

For a 1cm cube of ice, the enthalpy (heat) of melting is about 300 J, corresponding to 6 kJ mol^{-1} , and the resulting transition is extremely sharp. Contrast this with the melting of a (hypothetical) 20 Å cube of ice, about the same size as a protein molecule. Ignoring surface effects, this would require about 2.7x10^{-18} J (6.4x10^{-19} cal) to melt at 0 ^{O}C, or 1600 kJ mol^{-1} (380 kcal mol^{-1}), and would give the sigmoidal melting profile shown in Fig.3. Note that this is still a 2-state transition. There is no suggestion that the mini-ice cube is at any stage "half-

melted" - i.e. intermediate between liquid and solid. It simply shows that, for small systems, there is a finite range of temperatures over which significant populations of either state may be observed. The bigger the system, the more the cooperativity, the sharper the transition becomes until, for everyday macroscopic objects, the transition region is so narrow as to be imperceptible and the transition appear infinitely sharp.

This shows that, in the limit, even the most ideal, perfectly cooperative 2-state protein transition will have a finite width, determined solely by thermodynamic constraints.

2. Thermodynamics of Unfolding: Reversible Globular Proteins

2.1. Differential Scanning Calorimetry

Unfolding of proteins at elevated temperatures can be followed by a variety of indirect methods which, using the 2-state approximate analysis described above, can give information about thermodynamic parameters for the process. Much less ambiguous information, however, is given by calorimetric methods which measure energy changes directly. Differential scanning calorimetry (DSC), pioneered and developed for biomolecular studies by the Sturtevant, Brandts, and Privalov groups (Sturtevant, 1974, 1987; Jackson & Brandts, 1970; Privalov & Potekhin, 1986) is most applicable here. In a DSC experiment a solution of protein (typically 1 mg/ml or less in modern instruments) is heated at constant rate in the calorimeter cell alongside an identical reference cell containing buffer. Differences in heat energy uptake between the sample and reference cells required to maintain equal temperature correspond to differences in apparent heat capacity, and it is these differences in heat capacity that give direct information about the energetics of thermally-induced processes in the sample.

A typical DSC thermogram for the unfolding of a simple globular protein is shown in Figure 4.

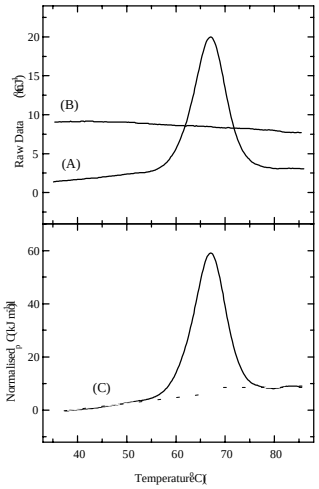

Fig.4: Typical DSC data for thermal unfolding of a globular protein. (A) Raw data - lysozyme, 3.7 mg/ml (0.26 mM), in 40mM glycine/HCl buffer, pH 3.0, scan rate 60 ^{O}C hr^{-1}. (B) Buffer baseline control, run under identical conditions. (C) Concentration normalised C_p data, with control baseline subtracted.

Note that, at most times, the heat capacity of the protein solution is lower than the control with buffer alone. This reflects the fact that protein, in common with most organic substances, has a lower heat capacity than liquid water. (Water is, of course, the unusual partner here, since the special features of its extended H-bonded structure endow it with a range of anomalous physical properties, including an unusually high heat capacity.)

After correction (by subtraction) of the buffer baseline control, three significant regions are apparent in this DSC trace. At low temperatures ("pre-transition") the heat capacity of the protein increases monotonically with temperature in a manner typical of organic solids. As the protein begins to unfold at higher temperatures the DSC trace becomes more positive, showing the increased apparent heat capacity arising from heat energy uptake in the endothermic unfolding transition. Once this transition is complete the thermogram reverts to a "post-transition" baseline, reflecting the heat capacity of the now-unfolded protein in solution. This post-transition baseline is characteristically off-set from the extrapolated pre-transition heat capacity, indicating a positive ΔC_p , and is usually flatter.

The shape and area of the transition endotherm contain thermodynamic information about the process. Most directly, the integrated area beneath the peak in the DSC endotherm, divided by the total amount of protein in the calorimeter cell, gives the calorimetric enthalpy (heat uptake, ΔH_{cal}) for the unfolding transition, independent of any model assumptions (apart from interpolation of pre- and post-transition baselines). Depending on how the protein concentration is measured, this might be quoted per mole or per gram of protein. The mid-point temperature of the transition (T_m) is the point at which 50% (on average) of the protein molecules are unfolded which, in simple cases, is the temperature at which the Gibbs free energy of unfolding (ΔG_{unf}) is zero.

Uniquely to DSC, a second and independent estimate of the unfolding enthalpy may be made from van't Hoff analysis of the shape of the peak in the DSC thermogram (Jackson & Brandts, 1970; Sturtevant, 1974,1987; Privalov & Khechinashvili, 1974; Hu *et al*., 1992). Assuming a 2-state transition model, the fractional heat uptake at any stage in the transition may be taken as a measure of the extent of unfolding and, as such, may be used just like any other (indirect) observable parameter to plot the fraction unfolded as a function of temperature. This fraction is an empirical quantity, independent of the sample concentration or absolute calorimetric enthalpy, and may be used as described earlier to estimate the van't Hoff enthalpy (ΔH_{VH}) of the process. This is the heat uptake per mole of cooperative unit in the transition, and comparison with the directly-determined calorimetric enthalpy (ΔH_{cal}) gives information about the size of the cooperative unit or the validity of the 2-state assumption. For an ideal, cooperative 2-state transition $\Delta H_{VH} = \Delta H_{cal}$, and this holds reasonably well (within 5%) for experiments involving small, simple globular proteins under conditions where their unfolding transition is reversible.

Frequently, however, this is not the case (Hu *et al*., 1992). Situations can arise where $\Delta H_{VH} > \Delta H_{cal}$, reflecting a DSC transition that is narrower than would be expected. This might indicate that the cooperative unit is greater than anticipated, due to specific dimer or higher oligomer formation for example, in which cases the ΔH_{VH} :ΔH_{cal} ratio is an indication of the number of protein molecules involved in the cooperative unfolding unit. Care must be exercised here, however, since anomalous sharpening or foreshortening of DSC peaks can (and frequently does) arise from irreversible processes such as exothermic aggregation of unfolded protein. Such effects can also have a kinetic component that will show up as a scan-rate dependence of the transition peak shape and position (Sanchez-Ruiz et al. 1988; Galisteo et al. 1991; Lepock et al., 1992).

The opposite situation, $\Delta H_{VH} < \Delta H_{cal}$, arises when the DSC transition is broader than would be expected for a 2-state transition with this particular ΔH_{cal} . This usually reflects a breakdown of the simple 2-state model assumption, indicating that unfolding of the protein involves several steps with at least one significantly populated intermediate phase. In some cases the thermogram might display clear shoulders or separate peaks that can be deconvoluted and correlated with the (possibly independent) unfolding of recognisable domains or subunits of the protein under investigation (Privalov, 1982).

2.2. Thermodynamics of Unfolding: Empirical Data

The DSC transitions of a range of small, monomeric globular proteins, including examples such as lysozyme, ribonuclease, myoglobin, cytochrome c, chymotrypsin and ubiquitin, have been extensively studied over the past 20-30 years as instrumental techniques have developed, and a consensus view is now appearing - at least for these relatively well-behaved proteins. Under most experimental conditions the thermal unfolding transitions of these proteins seem to follow cooperative 2-state behaviour well enough for us to ignore any significant build-up of intermediate states in the transition (Jackson & Brandts, 1970; Privalov, 1979). Calorimetric (ΔH_{cal}) and van't Hoff enthalpies (ΔH_{VH}) are close to identical within experimental error, i.e. ±5%, which is within the usual uncertainties associated with protein concentration measurements that are crucial to absolute molar ΔH_{cal} estimates. (Sometimes the ΔH_{VH} :ΔH_{cal} ratios are consistently slightly greater than one, possibly reflecting systematic errors in concentration measurement.)

As has been apparent for many years from a range of experimental methods, including DSC, in terms of thermodynamic free energy, folded proteins are only marginally stable with respect to their unfolded states. The experimental free energy difference (ΔG_{unf}) between folded and unfolded states under near-physiological conditions is usually in the range +20-60 kJ mol^{-1} (the positive sign reflecting the stability of the native fold). This corresponds to a stabilising free energy per amino acid residue much less than average thermal energy under these conditions ($k_BT \equiv 2.5$ kJ mol^{-1} at 300 K) and emphasises the cooperative nature of protein folding (Privalov, 1982, 1992; Murphy & Freire, 1992; Chan *et al*., 1995; for example): individually the interactions between amino acids are insufficient to maintain a stable conformation, but taken together in concert they are. For example, with a 100-residue protein an average value ΔG_{unf} = 40 kJ mol^{-1} corresponds to a 2-state equilibrium constant (K) of about 10^{-7} at 25 ^{O}C, implying that only one molecule in 10 million is cooperatively unfolded at any one time under these conditions. If, on the other hand, the polypeptide were able to "unravel" one or two residues at a time, the low free energy per residue (≈0.4 kJ mol^{-1}) would allow significant such unravelling. Presumably it is the strict topological or stereochemical constraints of the folded protein that usually do not allow such unconcerted actions - rather like a 3-dimensional jigsaw or "Chinese puzzle", where the removal of just one piece is impossible without disrupting the whole.

The temperature dependence of ΔG_{unf} shows that for most proteins the folded form is, not unreasonably perhaps, most stable in the physiological temperature range (see Figure 5). Variation of ΔG_{unf} with temperature is normally relatively small in the 20-40 ^{O}C region, but

with significant curvature as ΔG_{unf} falls to negative values at higher temperatures where the unfolded form becomes the more stable. The mid-point unfolding temperature (T_m) is given by the point at which this curve crosses the $\Delta G_{unf} = 0$ axis.

The relatively small free energy of unfolding is made up of, usually, much larger and much more temperature dependent enthalpy and entropy contributions. Unfolding is usually endothermic (but not always - see below), with a typical ΔH_{unf} of order +1 kJ mol^{-1} per residue at 25 ^{O}C, but varying rapidly and becoming increasingly more positive (more endothermic) with temperature. This positive ΔH_{unf} is offset by a (usually) positive entropic contribution, ΔS_{unf} , typically of order +2 J K^{-1} mol^{-1} per residue at 25 ^{O}C, but also increasing rapidly with temperature (Fig. 5).

This strong temperature dependence of ΔH_{unf} and ΔS_{unf} is a consequence of the heat capacity differences, ΔC_p , between folded and unfolded states. The heat capacity of the unfolded polypeptide chain, obtained by extrapolation of post-transition DSC baselines or from measurements on chemically unfolded samples (Privalov & Makhatadze, 1990), is higher than that of the folded protein (Fig. 4). For the unfolded protein the heat capacity appears to show relatively little variation with temperature, unlike the folded state where C_p generally increases with T (Jackson & Brandts, 1970; Brandts & Lin, 1990). As a consequence, ΔC_p itself also varies with temperature, becoming smaller at higher temperatures.

A word of caution regarding experimental ΔC_p estimates (Hu *et al*., 1992). Although in principle the ΔC_p for a protein unfolding transition may be obtained from the difference between extrapolated pre- and post-transition baselines in a single DSC experiment, in practice for most instruments the baselines are not well enough defined nor do they extend over a sufficient temperature range to assure confident extrapolation. Consequently an alternative experimental procedure is frequently adopted in which the T_m of the protein under investigation is varied in separate experiments, usually by variation of experimental pH. Analysis of the variation in ΔH_{unf} with T_m (essentially the slope of the ΔH_{unf} *versus* T_m plot) gives ΔC_p . In cases where comparison can be made, this approach gives ΔC_p values consistent with those measured directly from heat capacity baseline extrapolations, but it must be remembered that different transitions may be being observed under these differing experimental conditions and this might make additional contributions to ΔH_{unf} and, therefore, affect the apparent ΔC_p . Experiments done at different pH values, for example, will involve unfolding of differently ionised (charged) forms of the protein. It is unclear, at least at first sight, to what extent this will affect the measured heats or ΔC_p values. But comparison of the heats of unfolding of lysozyme at different temperatures by variation in both pH and denaturant (guanidinium chloride) concentrations (Privalov, 1979,1992; Pfeil & Privalov, 1976a,b,c) indicate that the unfolding enthalpy (for lysozyme, at least) is a function only of the temperature and not how the unfolding is brought about. Consequently, Privalov(1992) has argued that ΔC_p values determined in this way should be valid. However, the observation that ΔH_{unf} depends only on temperature and not on pH or

denaturant concentration is somewhat unexpected, and would imply that the reduction in stability of the folded protein by pH or denaturants is simply an entropic effect.

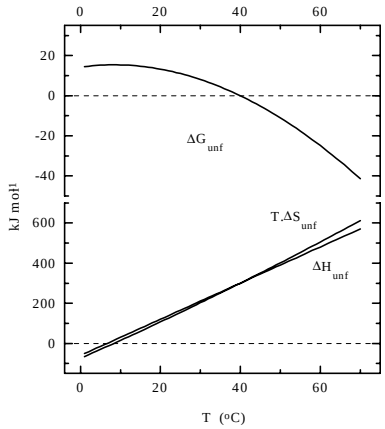

Fig.5: Characteristic temperature variation of thermodynamic parameters for unfolding of a small globular protein. Data are calculated for a typical protein unfolding at 40 ^{O}C (T_m) with ΔH_m = 300 kJ mol^{-1} and assuming a constant ΔC_p = 9 kJ K^{-1} mol^{-1} . Note how the relatively small unfolding free energy (ΔG_{unf}) is made up of the difference between relatively large enthalpic (ΔH_{unf}) and entropic (ΔS_{unf}) contributions. Temperature variation of ΔC_p would show as a curvature of the ΔH_{unf} and $T.\Delta S_{unf}$ lines.

2.3. Cold Denaturation

One significant consequence of a finite positive ΔC_p for the unfolding process is that the plot of ΔG_{unf} versus temperature is curved (Fig.5), decreasing either side of some intermediate temperature of maximum stability. At higher temperatures ΔG_{unf} eventually becomes negative, describing endothermic thermal unfolding (above). But similar extrapolation on the low temperature side suggests that, at some sufficiently low temperature, ΔG_{unf} should also be negative, suggesting that the unfolded protein should also become thermodynamically the more stable state at low temperature. This led to the prediction of exothermic “cold denaturation” of proteins (Brandts,1964; Franks, 1995) and was widely accepted as evidence for the dominant involvement of hydrophobic interactions in folding stability, since empirically the solubility of non-polar compounds in water is enhanced at lower temperatures. For most proteins under normal conditions the extrapolated temperature required for cold denaturation is below the freezing point of water, and different factors are expected to affect folding stability of proteins in a frozen matrix. But cold denaturation has been observed in a few instances, usually by addition of salts to depress the freezing point of the sample or by addition of denaturants that reduce the stability of the folded protein so that cold denaturation occurs at higher temperatures, above 0 ^{O}C. Calorimetric experiments on cold denaturation are technically quite difficult, but the limited amount of information gained so far suggests that cold denaturation behaves like a cooperative unfolding transition, with thermodynamic parameters consistent with estimated extrapolations from high temperature unfolding data (Privalov, 1990).

2.4. Thermodynamics of Unfolding: The Molecular Interpretation

Although the experimental situation regarding protein folding thermodynamics is now fairly well established, the interpretation of the thermodynamic parameters at the molecular level has been and remains much more controversial. Despite numerous reviews that have appeared in recent years, in addition to the classic Kauzmann (1959) article that first gave prominence to hydrophobic interactions, no clear picture has yet emerged. Particularly contentious has been interpretation of the temperature dependence of the unfolding enthalpies and entropies (ΔH_{unf} and ΔS_{unf}) where much has been made of the supposedly unusual or special "convergence" temperature(s) (usually in the region of 110 ^{O}C) at which extrapolated ΔS_{unf} and ΔH_{unf}, when expressed per mole of amino acid residue, were thought to achieve similar values for different proteins - (Privalov, 1979; Baldwin, 1986; Privalov & Gill, 1988; Murphy et al. 1990; Lee, 1991). It is now acknowledged that much of this speculation was based on over-interpretation of limited data from DSC experiments on a small set of proteins (Makhatadze & Privalov, 1995). More comprehensive analysis of accumulated more accurate data from an extended range of globular proteins allows a more rational overview.

Folding of a protein must overcome the thermodynamically unfavourable loss of conformational entropy associated with the dynamic heterogeneity of the conformationally disordered polypeptide in the unfolded state. Various estimates of this entropy have been made, both from theoretical considerations of the statistics of random coil polypeptides and extracted from experimental data (Schellman, 1955; Privalov, 1979; Brooks et al. 1988). Values range from 15 - 25 J K^{-1} mol^{-1} per residue arising from backbone conformational freedom (ϕ-ψ rotations, etc.), with additional contributions arising from restriction in side chain conformational mobility (Doig & Sternberg, 1995). This corresponds to a free energy ($T.\Delta S$) of order 6 kJ mol^{-1} or more per residue that must be overcome by a net negative contribution from changes in interactions between protein and solvent groups, either separately or collectively, in the folding process.

The fundamental problem in interpreting protein folding thermodynamics in terms of the individual molecular interactions between groups in the protein is, of course, that such interactions always involve differences between two states - typically the difference between a group exposed to solvent (water) in the unfolded protein, and buried in the folded form. It is the unavoidable involvement of solvent interactions, and particularly such a complex solvent as water, that makes analysis so difficult. Take, for example, the hydrogen bond interaction between two protein groups: the NH...O=C bond between peptide units, say. Such bonds are easily recognised in X-ray diffraction structures of proteins, and it is tempting to assume that they stabilise the structure. But, although H-bonds between buried peptide and other groups undisputedly stabilise the particular protein fold, it is even yet unclear to what extent they contribute to the overall stability with respect to the unfolded state. This is because in the unfolded protein the -NH and -C=O bonds (say) are presumably solvated (H-bonded to water molecules). During the folding process the H-bonds to water must be broken, then replaced by the intra-molecular bonds. Hence, in the overall process, taking solvent interactions into account as we must, there is no net gain in number of hydrogen

bonds in the system, though there will be entropic contributions arising from release of bound water that are less easy to visualise. Experiments with small model compounds seem to support this general picture (Klotz & Farnham, 1968; Kreshek & Klotz, 1969). Indeed, the ubiquitous high solubility of polar, H-bonding compounds in water shows that most groups "prefer" hydrogen bonding to water than to other groups - to the extent that model studies usually suggest that H-bonds within groups in proteins make an overall de-stabilizing contribution to the free energy of folding. That is, although it is energetically unfavourable to leave any H-bonds broken, it is relatively immaterial whether the H-bond is to a water molecule or to another protein group. So, when a protein folds, although all possible hydrogen bonds are probably made, their contribution to the folding free energy may be negligible or even repulsive. But other interpretations are possible (e.g. Dill, 1990a; Spolar et al., 1992). Similar problems afflict interpretation in terms of the other general kinds of interaction (electrostatic, hydrophobic, van der Waals) usually considered. (For general background introduction to forces, see: G.Allen - Vol.1, Ch.2 & references therein). Some aspects of electrostatic interactions are also considered further below.

Such considerations are usually based on analogies with small organic molecules in solid, liquid, vapour or solution states, and some success has been achieved in correlating thermodynamic parameters with changes in accessible surface areas of polar and non-polar groups on folding (Spolar et al., 1992). But the problem with small molecule model studies as analogues of the protein folding process is that such models rarely, if ever, mimic the detailed changes that occur between folded and unfolded proteins. And the importance may lie in the detail - to the extent that the best model systems may be the proteins themselves.

The complexities of the interpretative problem, and the ferocity of the arguments involved, are illustrated in two recent articles in the same volume of Advances in Protein Chemistry (Lazaridis *et al*., 1995; Makhatadze & Privalov, 1995; see, in particular, the epilogues to these chapters), as well as elsewhere (Makhatadze & Privalov, 1996).

Continuing a sequence of papers from this group, Makhatadze & Privalov (1995) present a detailed comparison of the published thermodynamic data from a range of proteins in comparison with their folded structures, and have attempted to dissect the interactions into their component parts to identify features characteristic of the different contributions. Their argument is too detailed to reproduce here, but in summary they conclude, somewhat surprisingly, that the dominant contribution to the stability of the compact folded protein comes from internal hydrogen bonding and, to a lesser extent, from the van der Waals attractions between closely packed groups within the protein. (Creighton, 1991, came to similar conclusions.) Little contribution appears to come from hydration of aliphatic groups, and burial of aromatic residues appears to be thermodynamically unfavourable, in contrast to received wisdom (Kauzmann, 1959; Dill, 1990). However, identification of the classic "hydrophobic effect" contribution within this scheme is difficult since Makhatadze & Privalov treat hydration and van der Waals contributions separately and in a way that makes comparison with other models less straightforward.

The numerical self-consistency of the Makhatadze & Privalov (1995) analysis is impressive. But it has to be said that the work is based on numbers extracted or extrapolated from published experimental data that appear, at least in some instances, to be more precise than the original raw experimental data or published figures would justify. It is also fair to say that equally convincing numerical correlations have appeared in the past based on similar data but with different parameters (for example: Makhatadze & Privalov, 1993; Privalov&

Makhatadze, 1992,1993; Khechinashvili et al., 1995). This probably reflects the sparsity of experimental data compared to the number of free parameters in any model.

In marked contrast, in the same volume of Advances in Protein Chemistry, Lazaridis *et al*. (1995), taking a different approach with a less extensive set of experimental data, come to markedly different conclusions regarding both the magnitudes and the signs of the different contributions to stability, supporting the more traditional view (Kauzmann, 1959; Dill, 1990a) that hydrophobic interactions are the primary source of folding stability and hydrogen bonding is the source of specificity of the folded conformation. They also point out some shortcomings in the Makhatadze & Privalov approach that might lead to overestimation of H-bonded contributions, for example. However, Laziridis *et al.* (1995) address only the enthalpic contributions to folding at one temperature, and it is not clear in this treatment where the important temperature dependence of enthalpies (ΔC_p) arises. Nor have they yet considered the much more difficult but equally important entropic terms.

Much of the disagreement between different models is often semantic, arising from different ways in which different workers elect to partition different contributions under different headings (Dill 1990b; Privalov et al. 1990, for example). Indeed, this desire to partition between different kinds of interaction may itself be flawed since, although it is understandable and would make contemplation of the problem easier, the various interactions are really in some ways just different manifestations of the same overall phenomena, and cannot necessarily be separated into individual, independent components. Hydrophobic interactions, for example, are just a manifestation of the hydrogen bonding properties of water. These same hydrogen bonds are responsible for the solvation of charged and polar groups that dominates the overall thermodynamics of H-bond formation in protein folding. Hydrogen bonds themselves are just a convenient construct: a way of visualising a particular sub-set from a larger class of polar interactions arising from permanent dipole/multipole effects. And all these interactions occur over a background of the unavoidable van der Waals interactions, with attractions arising from transient quantum mechanical charge fluctuations (London dispersion forces) and repulsions from too close approach of atoms. And it is probably a significant oversimplification to assume that all these interactions are necessarily additive, especially in such a cooperative structure as a folded protein.

Perhaps, too, we are asking rather too much at present when attempting detailed molecular interpretation of the empirical thermodynamic data. Even much simpler systems defy such analysis. The melting of a simple organic solid, for example, is not understood in the same detail that we seem to be demanding for protein unfolding. And the reason for this is instructive. Provided the crystal structure is known in sufficient detail, the intermolecular forces between small molecules in the solid can be computed relatively easily - this, in fact, forms the basis for many of the empirical force fields used in molecular mechanics calculations on proteins. But once the crystal melts we are in unknown territory. So little is known about the structure and dynamics of liquids at the molecular level that it is, as yet, impossible to calculate *ab initio* thermodynamic parameters (H, S, C_p) with sufficient confidence to estimate or even rationalise the crucial thermodynamic parameters (ΔH, ΔS, ΔC_p) for the melting phase transition. Compare this now with the protein folding situation. Even though it might be possible to obtain relatively good estimates of the energy of the folded polypeptide, it is the disordered, unfolded state which creates major difficulties. Not only do we have insufficient experimental data to characterise the population of conformational states that defines the unfolded protein, but each of these conformational states comprises a heterogeneous mixture of different molecular groups immersed in water,

which is a complicated enough molecular liquid in its own right. Interestingly, it is differences in assumptions regarding the nature of the unfolded polypeptide that lead, at least in part, to divergences in interpretation between Lazaridis *et al*. (1995), Makhatadze & Privalov (1995), and others. In such circumstances it is probably wise to regard with some circumspection detailed theoretical descriptions of the thermodynamic contributions to protein folding.

3. Effect of Ligand Binding on Folding Thermodynamics

Le Chatelier's principle implies that if any ligand (small molecule or other protein or macromolecule) binds preferentially to the folded protein, then this will stabilise the folded state and unfolding will become progressively less favourable as ligand concentration increases. Conversely, ligands that bind preferentially to the unfolded protein will destabilise the fold and will encourage unfolding. Examples of both are seen (Sturtevant, 1987; Fukada et al., 1983; Cooper, 1992; Cooper & McAuley-Hecht, 1993).

The general case of multiple ligands and multiple protein subunits has been considered by Sturtevant (Fukada et al. 1983; Sturtevant, 1987). For a simple case in which a ligand molecule (L) binds specifically only to the native folded protein (N), the following equilibria apply:

Ligand binding: $N + L \rightleftharpoons NL$; $K_{L,N} = [N][L]/[NL]$

Unfolding: $N \rightleftharpoons U$; $K_0 = [U]/[N]$

where $K_{L,N}$ is the dissociation constant for ligand binding to the native protein and K_0 is the unfolding equilibrium constant for the unliganded protein.

In the presence of ligand the effective unfolding equilibrium constant (K_{unf}) is given by the ratio of the total concentrations of unfolded to folded species:

$$K_{unf} = [U]/([N] + [NL]) = K_0/(1 + [L]/K_{L,N}) \approx K_0.K_{L,N}/[L]$$

where the approximate form holds at high free ligand concentrations ($[L] > K_{L,N}$). This shows that K_{unf} decreases and the folded form becomes more stable with increasing ligand concentration.

Expressed in free energy terms:

$$\Delta G_{unf} = -RT.\ln(K_{unf}) = \Delta G_{unf,0} + RT.\ln(1 + [L]/K_{L,N})$$

$$\approx \Delta G_{unf,0} + \Delta G^{O}_{diss,N} + RT.\ln[L] \quad \text{(for high [L])}$$

where $\Delta G_{unf,0}$ is the unfolding free energy of the unliganded protein, and $\Delta G^{O}_{diss,N} = -RT.\ln(K_{L,N})$ is the standard Gibbs free energy for dissociation of the ligand from its binding

site on the native protein. Thus the stabilising effect of bound ligand can be visualised as arising from the additional free energy required to remove the ligand prior to unfolding, together with an additional contribution (RT.ln[L]) from the entropy of mixing of the freed ligand with the bulk solvent.

In the high ligand concentration limit the free energy can be separated into enthalpy and entropy contributions thus:

$$\Delta H_{unf} \approx \Delta H_{unf,0} + \Delta H^{O}_{diss,N}$$

and $\Delta S_{unf} \approx \Delta S_{unf,0} + \Delta S^{O}_{diss,N} - R.\ln[L]$

For small ligands the heat of dissociation ($\Delta H^{O}_{diss,N}$) can be quite small compared to the heat of unfolding of the protein, and may be hard to distinguish in calorimetric unfolding experiments, particularly when ΔH_{unf} is in any case varying with temperature due to ΔC_p effects. Entropy effects, particularly those arising from the ligand mixing term (R.ln[L]), will be much more apparent in such cases.

[Slightly more complex, but manageable expressions, corrected for the fraction of unliganded protein in the mixture, apply at lower concentrations of ligand. In such cases the thermodynamic parameters have values intermediate between unliganded and fully-liganded values given above.]

Similar considerations apply in situations where ligand binds only to the *unfolded* protein (Cooper, 1992; Cooper & McAuley-Hecht, 1993):

$$U + L \rightleftharpoons UL \quad ; \quad K_{L,U} = [U][L]/[UL]$$

in which case:

$$K_{unf} = ([U] + [UL])/[N] = K_0.(1 + [L]/K_{L,U}) \approx K_0.[L]/K_{L,U}$$

and: $\Delta G_{unf} = -RT.\ln(K_{unf}) = \Delta G_{unf,0} - RT.\ln(1 + [L]/K_{L,U})$

$$\approx \Delta G_{unf,0} - \Delta G^{O}_{diss,U} - RT.\ln[L] \qquad \text{(for high [L])}$$

which in this case shows the destabilising effect of a reduction in unfolding free energy as ligand binds to the unfolded polypeptide. Equivalent expressions for the enthalpy and entropy contributions may be written as above, with appropriate sign changes.

An example of this kind of effect is illustrated in Fig. 6 for the unfolding of globular proteins in the presence of cyclodextrins. These toroidal oligosaccharide molecules form inclusion complexes with small non-polar molecules and therefore bind to exposed aromatic groups on the unfolded protein (Cooper, 1992).

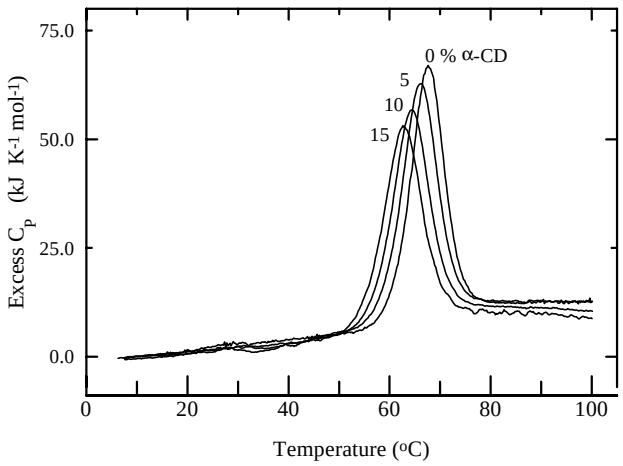

Fig.6: DSC traces showing the effect of increasing α-cyclodextrin concentrations (0-15% w/v) on the thermal unfolding of lysozyme (40mM glycine buffer, pH 3.0). Note the progressive reduction in both T_m and apparent ΔC_p .

Note: the apparent variation in ΔH_{cal} is predominantly due to the inherent variation of unfolding enthalpy with temperature (ΔC_p effect) rather than the result of ligand binding *per se*.

The effect of ligand binding (either to N or U) on T_m of the protein can be generalised and approximated in the case of weakly binding ligands (Cooper & McAuley-Hecht, 1993) to give:

$$\Delta T_m / T_m = \pm(nRT_{m0} / \Delta H_{unf.0}).\ln(1 + [L]/K_L)$$

where $\Delta T_m = T_m - T_{m0}$ is the shift in unfolding transition temperature and n is the number of ligand binding sites on the protein (assumed identical). The ± sign relates to whether ligand stabilises the folded or unfolded form.

At low concentrations, with weakly binding ligands ([L]/KL << 1) this becomes approximately linear in ligand concentration:

$$\Delta T_m / T_m \approx \pm nRT_{m0} [L]/(K_L.\Delta H_{unf.0})$$

Note that the T_m shift continues with increasing ligand concentration even beyond levels where the protein is fully ligand-bound. This is a manifestation of the dominant entropy of mixing contribution described above. Cases do arise, however, where the T_m shift does plateau at higher ligand concentrations. This usually signifies binding of L to *both* N and U, albeit with different affinities. For example, a particular ligand might bind strongly to the native protein but less well to the unfolded chain. In such cases the T_m would shift upwards with increasing [L] until the concentration is such that both N and U are fully liganded. A recent example of this is α-lactalbumin (Robertson, Cooper & Creighton - in preparation), a specific calcium binding protein where increasing [Ca^{2+}] increasingly stabilises the native protein up to a limit where weak, non-specific calcium ion binding to the unfolded chain sets in.

Analysis of more complex situations involving multiple ligand binding or more tightly binding ligands is generally less straightforward, but the same basic principles apply. See Sturtevant (1987), Brandts & Lin (1990) for details.

3.1. Effect of pH

The effect of varying pH on the stability of protein folding is just a special case of the ligand-binding consequences described above. In this case the ligands are aqueous hydrogen ions (H^+) that bind to specific protein sites (acidic or basic groups) in both folded and unfolded states. Only if the proton binding affinities differ between the two states will pH have any effect on stability.

Consider, for simplicity, the proton binding to a single group on the polypeptide. The acid-base equilibrium for folded and unfolded states may be described:

$$N + H^+ \rightleftharpoons NH^+ \quad ; \quad K_{A,N} = [N][H^+]/[NH^+]$$

$$U + H^+ \rightleftharpoons UH^+ \quad ; \quad K_{A,U} = [U][H^+]/[UH^+]$$

The apparent or effective equilibrium constant for protein unfolding in this case is given by:

$$K_{unf} = ([U] + [UH^+])/([N] + [NH^+]) = K_0.(1 + [H^+]/K_{A,U})/ (1 + [H^+]/K_{A,N})$$

where $K_0 = [U]/[N]$ is the unfolding equilibrium constant for the unprotonated species.

It follows from this that the stability of the folded protein (with respect to unfolded) can only be affected by changes in pH if $K_{A,N} \neq K_{A,U}$.

The pH-dependence in more realistic situations with multiple ionisable groups is somewhat more complex, but the general principle still applies that changes in pH can only affect folding stability if the ionisable group(s) have different pK_A values in the folded and unfolded states.

It also follows from the above that, in regions where the stability of the folded protein is sensitive to pH, the folding <---> unfolding transition must be accompanied by an uptake or release of hydrogen ions. Using the general theory of linked thermodynamic functions (Wyman, 1964; Wyman & Gill, 1990), the mean change in number of H^+ ions bound when the protein unfolds is given by:

$$\delta n_{H+} = - \partial \log K_{unf} / \partial pH$$

Shifts in pK (δpK) correspond to changes in standard free energy of proton ionisation of the group ($\delta\Delta G^O_{ion}$) which are numerically related by:

$$\delta\Delta G^O_{ion} = -2.303RT.\delta pK$$

where R is the universal gas constant (8.314 J K^{-1} mol^{-1}) and T the absolute temperature. This corresponds to almost 6 kJ mol^{-1} per unit shift in pK at room temperature.

Figure 7 gives an illustration of the effect of pH on the thermal unfolding of a simple globular protein (lysozyme) as seen in DSC experiments. The major change in T_m in the low-pH region occurs over the pH 2-3 range, consistent with protonation of carboxylate side chains, and the variation corresponds to a maximal uptake (δn_{H+}) of about 3 hydrogen ions during unfolding of this protein under these conditions. This does not, of course, imply that there are three specific titrating groups responsible for this behaviour, but rather that this is the cumulative effect of all participating groups.

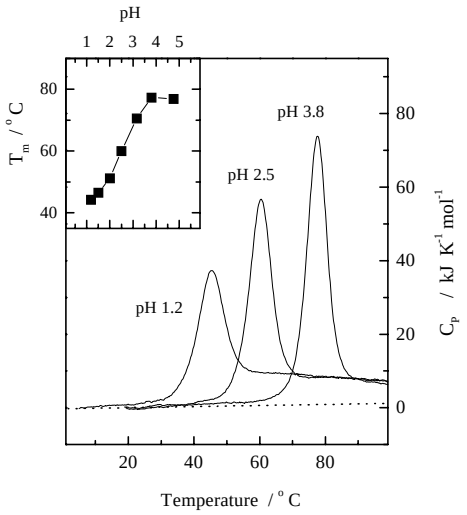

Fig.7: Effect of pH on the thermal unfolding of lysozyme in the DSC. The insert shows the variation in T_m with pH for this protein.

3.2. Electrostatic Interactions

Changes in group pK_A can be brought about by variations in effective polarity (dielectric constant) of the environment as a result of burial of residues within the folded protein for example, or by electrostatic interactions with other charged groups (Stigter & Dill, 1990; Yang et al., 1993; Antosiewicz et al., 1994; and references therein). All these factors are likely to change when a protein unfolds, so it is not unexpected that pK_A's might be different between the two states. In some cases the pK_A shifts can be quite large, 3-4 pK units for example in specific instances involving short-range electrostatic interactions or burial in non-polar locations, usually for residues with important catalytic or other specific functions. But generally the pK shifts for most residues are much smaller than this, since most charged groups are usually found close to the outer surface of the folded protein, and only relatively small changes in electrical environment occur on unfolding. Nevertheless, the accumulation of small pK_A shifts from a large number of such groups will make a considerable contribution, and the folding stability of most proteins is therefore sensitive to pH.

Exact calculation of electrostatic properties of proteins is a complex and computationally intensive problem (Stigter & Dill, 1990; Yang et al., 1993; Antosiewicz et al., 1994). But simple Coulomb interaction models can give an interesting and, perhaps, somewhat unexpected view of the complexity of the thermodynamics of charged groups in proteins. For

example, assuming point charges and a uniform dielectric medium, the electrostatic free energy (δG_{el}) between two charges, q_1 and q_2, with a distance R_{12} between them, is given by the classic Coulomb energy: $q_1q_2/4\pi\varepsilon_0\varepsilon R_{12}$, where ε_0 is the permittivity of free space and ε is the relative dielectric constant of the medium around the charges. This can be viewed as the work done, thus free energy change, in bringing these charges together from infinity to a separation R_{12}. For singly-charged groups and with R_{12} expressed in Angstrom units (Å) this can be written:

$$\delta G_{el} = \pm 1380/\varepsilon R_{12} \quad \text{kJ mol}^{-1}$$

where the ± sign depends on whether interactions are attractive (opposite charges, negative δG_{el}) or repulsive (like charges, positive δG_{el}). For charged groups separated by, say, 5 Å this amounts to about 3.5 kJ mol^{-1} in water at 25 ^{O}C with a dielectric constant of about 80, and corresponds to a combined pK shift of about 0.6 pK units. However, in a much lower dielectric environment such as the interior of a protein ($\varepsilon \approx 2.5$ to 4; Gilson & Honig, 1986) this can rise to $\delta G_{el} \approx$ 100 kJ mol^{-1} and (probably unrealistically) a combined δpK in excess of 12.

Burial of individual charged groups within the non-polar environment of a folded protein is generally energetically unfavourable. Again assuming a continuous dielectric, the free energy of transfer of a single spherical charge (q) of radius r from medium 1 to medium 2 is given by:

$$\delta G_{trans} = q^2(1/\varepsilon_2 - 1/\varepsilon_1)/8\pi\varepsilon_0 r \equiv 690(1/\varepsilon_2 - 1/\varepsilon_1)/r \quad \text{kJ mol}^{-1}$$

with r in Å for a single charge in the latter case. Taking a representative atomic radius ($r \approx 2$ Å) with $\varepsilon_1 = 80$ and $\varepsilon_2 = 4$ that might be a typical for burial of a group within a protein, this gives $\delta G_{trans} \approx 80$ kJ mol^{-1} .

Calculations such as these are simplistic: the continuum dielectric model is unrealistic at the atomic level, and we have ignored screening and other effects due to buffer electrolytes, for example. Nevertheless, they do illustrate the potential importance of charge interactions to folding stability, and these are the sorts of numbers that come out of more rigorous calculations and from experiment (e.g.: Dao-pin et al., 1991).

The partitioning of these electrostatic free energies into enthalpy and entropy components is also complicated. For any given geometry, the temperature dependence of the electrostatic free energies will depend on the temperature dependence of ε. Interestingly, since dielectric constants generally decrease with increasing temperature, at least in fluid environments, this means that an electrostatic attraction between two groups actually gets stronger, in free energy terms, the higher the temperature. Thermodynamically this would imply a positive ΔS contribution to the attractive free energy between oppositely charged groups. This can be rationalised in terms of the dipole-orientation entropy of molecules in the dielectric medium. Model studies of electrostatic interactions in salts or solutions bear out the complexity of the thermodynamics of such interactions, which may be endothermic or exothermic, entropy-driven or not, as the case may be.

The complex electrostatic properties of real proteins have received detailed attention only relatively recently (Gilson & Honig, 1986; Stigter & Dill, 1990; Yang et al., 1993; Antosiewicz et al., 1994), and the breakdown into enthalpy/entropy contributions is still unclear.

3.3. Denaturant and Osmolytes

There is still considerable discussion regarding the mechanism of unfolding of proteins by chemical denaturants such as urea, guanidinium chloride, etc. Possibly the effect arises from (weak) binding of these molecules to groups on the unfolded protein that would destabilise the folded form in the manner described above for other ligand-binding situations (Makhatadze & Privalov, 1992). Alternatively it is suggested that the effect is more indirect, resulting from changes in solvent structure or hydration/solvation of the protein, especially at the high concentrations at which these chemical denaturants are effective (Schellman, 1987a,b; Timasheff, 1992). Nevertheless, regardless of the detailed mechanism, denaturation by high concentrations of urea, guanidine chloride, or other highly water soluble compounds has long been recognised as a useful empirical tool. It is widely used in studies of site-directed mutagenesis effects on protein stability (e.g.: Matouschek et al., 1994; Serrano et al., 1992; Fersht et al., 1992) where it has been particularly effective in estimating the small changes (usually) in folding free energy brought about by amino acid replacements or other minor modifications. The procedure is based on extrapolation of free energy and other data obtained over a range of denaturant concentrations. Typically the extents of unfolding at different urea or GuHCl concentrations might be measured by CD, fluorescence, or other technique, and converted to a ΔG_{unf} using a 2-state assumption, as described earlier. These data correspond, of course, to unfolding free energies at relatively high denaturant concentrations (e.g. 2-8 M) and are not necessarily related to more physiological conditions. Empirically, however, it is found that ΔG_{unf} varies almost linearly with denaturant concentration and can be extrapolation to zero concentration to give an estimate in the absence of denaturant. This extrapolation is quite long, and concern has been expressed about its validity, but detailed comparisons of this method with more direct calorimetric determinations show remarkably good agreement (Hu *et al*., 1992; Santoro & Bolen, 1992; Matouschek et al., 1994; Johnson & Fersht, 1995), though the extrapolations are not always linear and care has to be taken to maintain a sufficiently high salt concentration in the case of GuHCl denaturation.

Addition of alcohols and other miscible solvents also generally reduces the stability of proteins in water. The thermodynamics of this (Velicelebi & Sturtevant, 1979; Woolfson et al., 1993) are consistent with what might be expected from reduction in hydrophobic interactions resulting from reduced polarity of the solvent environment of the unfolded polypeptide. But detailed analysis is complicated because of the inevitable effect such drastic solvent changes will have on the conformational population of the unfolded chain, which is even less likely to be “random coil” in the presence of organic solvent mixtures.

Osmolytes, on the other hand, are a range of water-soluble compounds that, at relatively high concentrations and in contrast to denaturants, stabilise globular proteins against thermal unfolding (Santoro et al., 1992). Such effects are biologically important in organisms subjected to heat, dehydration or other environmental stress, where a range of naturally-occuring osmolytes including sugars, polyhydric alcohols, amino acids and methylamines

may protect against protein denaturation (Yancey *et al*., 1982). Glycine based osmolytes such as sarcosine (8.2M concentration) give an increase in T_m of up to 23 ^{O}C, for example, with small globular proteins (Santoro et al., 1992). The mechanism of osmolyte stabilization of folded proteins remains unclear.

4. "Molten Globules" and other non-native states

The thermodynamic properties of molten globules and other non-native protein conformations are difficult to establish unambiguously (Privalov, 1996). Partly this is because the states themselves are difficult to define, and in only relatively few instances can experimental conditions be found which stabilise significant populations of such species. Also, by their very nature, such states lack the cooperativity characteristic of folding to the compact native conformation. This means that the 2-state model is rarely applicable to transitions to or from the molten globule state. Instead, changes in temperature or other experimental variable usually give rise to continuous changes in properties consistent with a more gradual shift in conformational population. In such situations only calorimetric methods can give unambiguous thermodynamic data, and even here the data are sparse. DSC experiments on the thermal unfolding of the "acid molten globule state" of apo-myoglobin for example (Griko & Privalov, 1994; Makhatadze &Privalov, 1995) show only a gradual heat energy uptake and a broad, sigmoidal increase in heat capacity with temperature, with none of the cooperative endothermic heat capacity discontinuity seen for the true native protein at higher pH. Similar results are found with α-lactalbumin (Griko et al., 1994) and other proteins, though comparative discussion is often hampered by lack of agreed definition and characterization of these states. In such a situation it is fair to ask whether the molten globule is really such a well defined state. Ptitsyn (1995; and earlier references therein) has argued strongly that it is. But the lack of any well defined thermal transition suggests the more general view that we are seeing just variation in a continuum of conformationally heterogeneous states under conditions where the native fold is only marginally stable. Observation of molten globule states typically requires low pH (pH 2-4), lack of co-factor or ligand (e.g. apo-α-lactalbumin lacking bound Ca^{2+} ; apo-myoglobin lacking the heme group), sometimes with addition of low concentrations of denaturant (alcohols, GuHCl, etc.). Under such conditions the protonation of acidic residues and the lack of stabilising ligand interactions will tend to destabilise the native fold. Yet, particularly at low temperatures, there will be sufficient residual interactions between residues to support clustering of conformations in more compact states, possibly even resembling the native state in secondary structure content and other properties (Griko et al., 1994). But with increase in temperature or harsher pH/denaturant conditions, the conformational heterogeneity will gradually expand to more open states, spanning greater regions of conformational space. In such a broad continuum of conformationally heterogeneous states it is a matter of taste or experimental convenience where one draws the line between "native", "molten globule", "partially folded", or "unfolded" states. Moreover, different experimental techniques will probe different aspects of these conformational populations and may give conflicting views. See Privalov (1996) for a critical review.

5. Reversibility

Central to all of the thermodynamic discussion, and to most experimental determinations of thermodynamic parameters for folding transitions, is the assumption that the process under investigation is reversible - that is, on the time scale of the experiment, that the system is in equilibrium and the concentrations of all molecular species present is determined by thermodynamics and not kinetics. This is frequently not the case, and can be a particular problem with experiments involving thermal unfolding (DSC for example) where exposure of the unfolded polypeptide to relatively high temperatures can bring about a variety of physical and chemical changes that affect the reversibility of the folding and can prejudice the results unless carefully controlled. Chemical changes such as proline isomerization, disulphide interchange, oxidation, and spontaneous de-amidation of asn and gln residues, for example, are all possible and will alter the folding properties of the polypeptide. Aggregation or precipitation of the unfolded polypeptide is also common at high temperatures or in certain solvent mixtures.

In calorimetric experiments, such irreversible processes can be recognised by their effects of the thermogram. Fig.8 for example, shows a series of repeated DSC traces for the thermal unfolding of lysozyme, where the sample is simply cooled back to room temperature after each scan. Although the major, native transition at about $74^{O}C$ is apparent throughout, each successive heating/cooling cycle sees the appearance of two (or more) transitions at lower temperatures together with a decrease in magnitude of the main transition. These less stable species are probably mis-folded, or incorrectly folded forms of the polypeptide brought about by the build up of chemical changes (proline isomerization, side chain de-amidation) with repeated unfolding and exposure to high temperature (Cooper & Nutley, unpublished). Although proline isomerization is reversible, in principle (Stein, 1993; Schmid et al., 1993), it is likely to be slow on the timescale of these experiments, such that on cooling the polypeptide gets trapped with the wrong proline conformers. Lysozyme has two proline residues in its amino acid sequence, so four different cis/trans combinations are possible in principle - though both are trans in the native conformation. It is interesting, but by no means yet conclusive, to note the appearance of 4 possible misfolded species in the DSC experiment (Fig. 8). Disulphide effects are not likely here since the process appears unaffected by the presence of reducing agents (DTT).

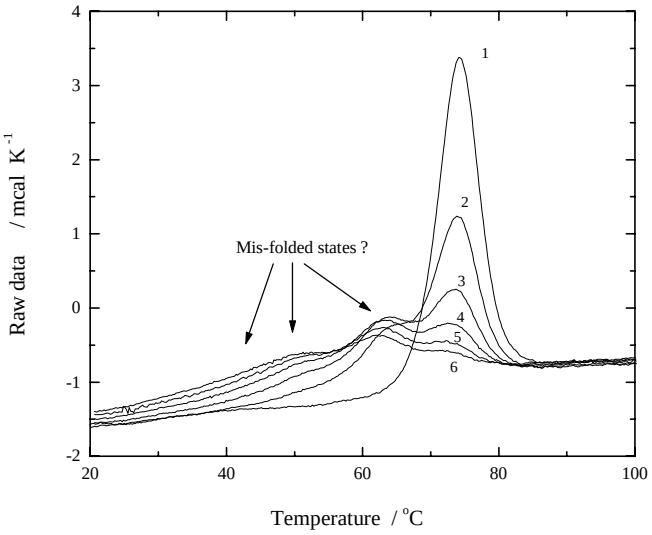

Fig.8: Repeat DSC scans of thermal unfolding of lysozyme (3.12 mg/ml, 0.1M glycine/HCl, pH 3.4) showing possible accumulation of misfolded forms. Scan rate was 60 ^{O}C hr^{-1}, with 60 min. cooling between scans.

This contrasts with another example where we have shown that a time-dependent irreversible effect on the folding of the methionine repressor protein, MetJ, can be totally eliminated by addition of DTT to the sample buffer (Johnson *et al*., 1992). Figure 9 shows a series of repeat DSC scans of MetJ giving a progressive decrease in magnitude with each heat/cool cycle. No misfolded species are apparent here, nor is there any evidence of thermal aggregation of the protein, but the effect depends on the amount of time the polypeptide is kept in the unfolded state at high temperatures and appears to be related to disulphide exchange, since it can be suppressed by addition of DTT. In the absence of reducing agents the kinetics of loss of refolding capacity are roughly first order in time above the unfolding temperature (Fig. 9).

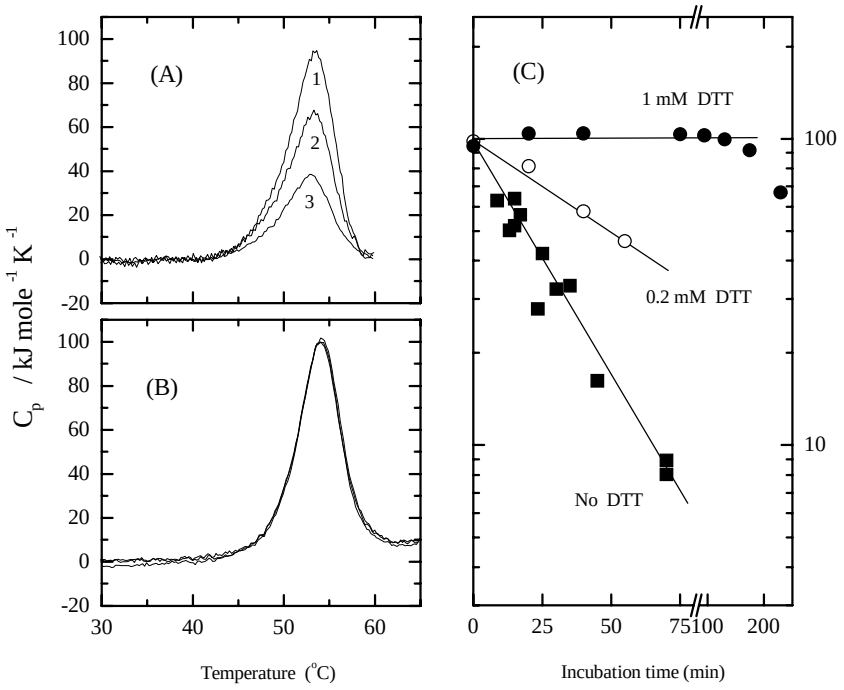

Fig.9: Effect of reducing agent on the reversibility of thermal unfolding of the methionine repressor protein (MetJ). (A) Repeat DSC scans of MetJ in the absence of reducing agent. (B) Repeat DSC scans of MetJ in the presence of 1mM DTT (dithiothreitol). (C) Effect of DTT on the degree of reversibility of the MetJ thermal unfolding transition following different incubation periods above 45 ^{O}C (see Cooper et al., 1992, for details).

In the case of MetJ explanation of this effect is relatively straightforward. MetJ is a dimeric protein, and each monomer contains one buried cysteine (-SH) residue whose function is (as yet) unknown, but which remains reduced in the native dimer structure. Upon unfolding, under oxidizing conditions the formation of intermolecular S-S crosslinks between these cysteines is likely, giving non-native crosslinked dimers that are unable to fold correctly. [It is tempting to speculate that such non-native, crosslinked dimers might actually be transient intermediates in the protein folding pathway of this dimeric protein in the reducing conditions found within the cell, since this would facilitate correct juxtaposition of the monomers prior to folding, but this hypothesis has yet to be tested.]

For the lysozyme and MetJ examples quoted above, the irreversible processes are usually too slow to have any serious effect of the DSC measurements, or can be eliminated by addition of appropriate reducing agent. Frequently, however, this is not the case, and serious distortion of DSC thermograms results from (usually exothermic) irreversible processes occurring simultaneously with thermal unfolding. Thermal aggregation (precipitation) of unfolded protein is a particular problem. This is illustrated (for PGK) in Fig.10 where the shape of the thermogram is severely distorted by exothermic aggregation of the unfolded polypeptide, and the noisy post-transition baseline is a consequence of erratic convection effects of precipitated protein within the calorimeter cell.

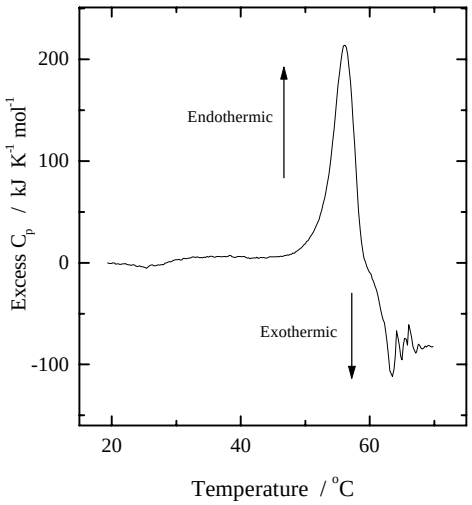

Fig.10: DSC data for thermal denaturation of yeast phosphoglycerate kinase (PGK; 50mM Pipes, pH 7.0) illustrating exothermic baseline distortion and noise effects caused by irreversible precipitation of unfolded protein.

Such aggregation is rarely reversible, and rescan of such samples after cooling show no discernible transition.

Even when no irreversible effects are immediately apparent from the shape of the DSC thermogram, a dependence on DSC scan rate can often indicate problems. Several groups (Sanchez-Ruiz et al. 1988; Galisteo et al. 1991; Lepock et al., 1992) have done detailed analysis of such situations and have developed theoretical procedures that allow such experiments to give both thermodynamic and kinetic information.

Irreversibility (or non-reversibility) is also apparent in many non-calorimetric experiments, where it can be monitored by lack of total regain of enzyme activity, for example, or simple appearance of protein precipitate (see discussion by Mitraki et al., 1987, for example). The possible distortion that such effects may produce on equilibrium denaturation curves has been less systematically explored, as yet.

6. Effects of Crosslinking

The presence of irreversible crosslinks, in the form of -S-S- bridges between cysteine residues or other covalent links connecting regions of polypeptide, enhances the relative stability of the folded protein, and the introduction of such crosslinks is a very effective way of improving stability. The effect is primarily entropic, arising from the topological constraints leading to a reduction in the number of configurations available to the unfolded chain (Schellman, 1955; Flory, 1956; Poland & Scheraga, 1965; Pace et al., 1988). In the absence of crosslinks, the distance between any two groups in the unfolded protein varies, with a probability distribution determined by the statistics of the polymer chain and a range dependent only on the length of the chain. A crosslink between two distant groups in the polymer forms a loop with a much restricted set of possible chain configurations, with a statistical distribution restricted to only those conformers that give an end-to-end chain distance consistent the juxtaposition of groups enforced by the crosslinks.

For any one loop, formed by crosslinking between groups n residues apart in the chain sequence, using classical theories of polymer chain statistics (Jacobson & Stockmayer, 1950; Schellman, 1955; Flory, 1956), the reduction in conformational entropy (ΔS_{conf}) of the

unfolded chain estimated by considering the relative probability that the ends of a polymer chain will be found within the same volume element (v_s) is given by:

$$\Delta S_{conf} = -R.\ln(3/(2\pi l^2 n)^{3/2})v_s$$

where l is the length of a statistical segment of the chain, usually taken to be 3.8 Å for a polypeptide. (This is for a single loop. The more complex situation of multiple, topologically dependent loops has been considered by Poland and Scheraga, 1965).

Various estimates of vs have been used. For a disulphide crossbridge, taking the distance of closest approach of the -SH groups as about 4.8 Å (Thornton, 1981), Pace et al. (1988) used $v_s = 57.9$ Å^3 (corresponding to a sphere of diameter 4.8 Å) giving:

$$\Delta S_{conf} = -8.8 - (3/2)R.\ln(n) \qquad J\,K^{-1}\,mol^{-1}$$

which gave reasonable agreement with experiment for the decrease in folding free energy ($\delta\Delta G = T.\Delta S_{conf}$) of a series of proteins upon removal of specific disulphide bridges.

Such agreement may be fortuitous however, since there are various assumptions and approximations inherent in the above. In particular, it is assumed that the unfolded polypeptide behaves as a statistical random coil, with a Gaussian end-to-end chain probability distribution in the absence of crosslinks. This may be reasonable for relatively large loops in a good denaturing solvent mixture, but will probably overestimate the effect under more realistic situations with most proteins, where the experimentally accessible unfolded state probably still contains residual conformation and is less than random coil. Furthermore, these estimates assume that the crosslink effect lies simply in the configurational entropy of the unfolded chain, and that the presence of the crosslink in the folded protein introduces no conformational strain or other constraints in the native form. Doig and Williams (1991) have also argued that the presence of disulphide crosslinks in the unfolded polypeptide leads to strain and other additional effects in the unfolded protein that override the entropic effects, though earlier work appears to rule this out (Johnson et al., 1978).

These various possibilities have been explored by more detailed thermodynamic measurements of specifically disulphide modified proteins (Cooper et al., 1992; Kuroki et al., 1992), with somewhat divergent conclusions, though care must be exercised to ensure that the experimental modifications used do not introduce additional destabilizing effects into the folded protein in the form of bulky or charged substituents.

DSC comparison of the thermal unfolding of native (4-disulphide) and a specific 3-disulphide hen egg white lysozyme is illustrated in Fig.11 (Cooper et al., 1992). Removal of the Cys6-Cys127 crossbridge results in a reduction in T_m of 25 to 30 ^{O}C under the same conditions together with a reduction in ΔH_m. However, because of the inherent variation of ΔH_m with temperature (ΔC_p effect) it is not possible from one such experiment alone to determine the source of destabilization. Comparison of ΔH_m for these proteins over a range of temperatures (by conducting experiments over a range of pH) shows that, within experimental uncertainty, the enthalpies of unfolding of these two proteins fall on the same line and that, for unfolding at the same temperature, the enthalpies are the same.

Consequently any difference in folding stability must arise solely from entropy differences. ΔS_{unf} for the two proteins (Fig.12) differ by about 90 J K^{-1} mol^{-1} over the pH range studied, in reasonable agreement with theoretical estimates for a 122 residue loop (Pace et al., 1988).

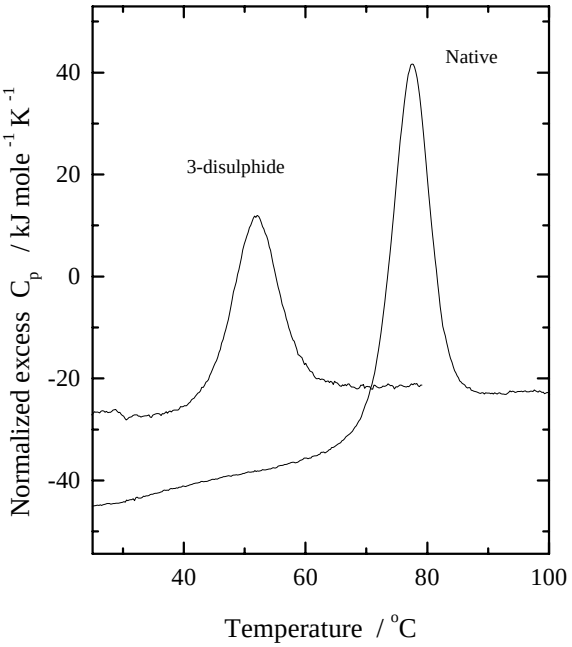

Fig.11: DSC comparison of thermal unfolding of native (4-disulphide) and $CM^{6,127}$ (3-disulphide) lysozyme at pH 3.8, 50mM glycine/HCl buffer.

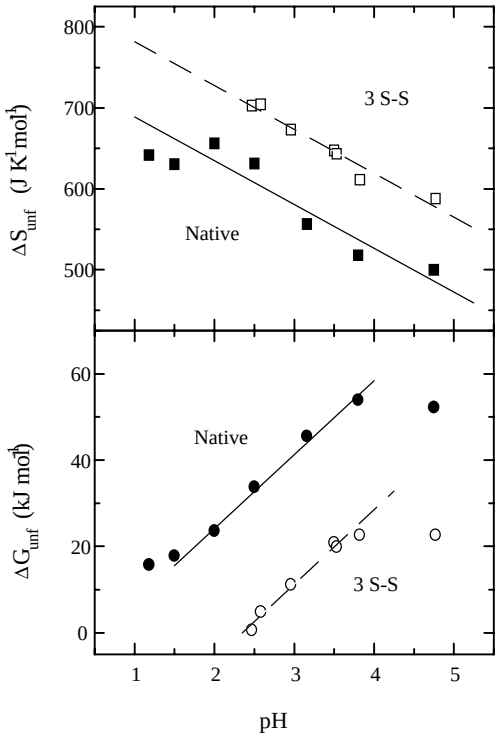

Fig.12: Variation with pH of the entropy (ΔS_{unf} , upper panel) and free energy (ΔG_{unf} , lower panel) of unfolding at 25 ^{O}C of native lysozyme and its 3-disulphide form. The lines in the lower panel show the free energy behaviour expected for an uptake of 3 H^{+} ions per molecule during unfolding.

The disulfide modification used here and the location of this particular crossbridge in the native structure is such that minimal perturbation of the folded protein is expected here, and this is confirmed by NMR studies (Radford et al., 1991).

Qualitatively similar results have been obtained in a recent comparison of the thermal unfolding of native bovine α-lactalbumin and a modified form lacking the equivalent 6-120 disulfide bond (Robertson, Creighton & Cooper, unpublished/in preparation - 1996). Here, however, although the enthalpies of unfolding of the two forms of the protein are similar when compared at the same temperature, the destabilizing effect and the entropy difference is somewhat less than would be anticipated for a loop of this size using the theory above. There are various possibilities for this discrepancy. Firstly, "unfolded" α-lactalbumin is known to exist in a range of different conformational sub-classes (including "molten globule") depending on conditions, and it is unlikely to behave as a fully random coil upon thermal unfolding. The system is yet more complicated by the Ca^{2+} binding of this protein, and Ca^{2+} or other cation binding to the unfolded polypeptide might produce transient non-covalent crossbridges and further restrict the conformational freedom of the chain. Moreover, tryptophan fluorescence quenching experiments (unpublished) of the folded protein indicate that removal of this disulfide link increases the accessibility of some trp residues to small molecule quenchers, thus indicating that the conformation or conformational dynamics of the native form seem also to be affected by removal of this crossbridge. No NMR or crystallographic data are yet available to check this more thoroughly.

By contrast, studies by Kuroki et al. (1992) of mutant human lysozymes lacking the disulfide bridge between cysteine residues 77 and 95 indicate that the observed destabilization in this case is enthalpic, with a paradoxically smaller unfolding entropy for the mutants lacking this crosslink. The difference here may be because the Cys77-95 crosslink involves a relatively tight loop and is buried within the protein structure rather than close to the surface as in the previous examples. Consequently, removal of this crossbridge is likely to have significantly greater effect on the native structure and dynamics. Kuroki et al. (1992) indeed showed that removal of this link did increase the flexibility of the native state, thereby increasing the entropy of the folded form of the protein. More recent studies on another protein (Vogl et al., 1995) confirm this general trend that relative contributions to folding stability of enthalpic and entropic terms depends on loop length and positioning of the crossbridge. Destabilization involving large loops tends to be purely entropic, as expected from the classic picture, but enthalpy effects play a greater role for shorter loops.

7. Fibrous Proteins

Relatively little systematic work has been done on the thermodynamics of folding of fibrous or other non-globular proteins. Experimentally such proteins are frequently more difficult to work with. They are often poorly soluble and difficult to purify to homogeneity in sufficient quantities for biophysical studies. They are generally high molecular weight, made up of several long polypeptide chains that makes them prone to aggregation and entanglement when unfolded. The unfolding transitions are therefore often irreversible on the experimental timescale, and non-cooperative or non-2-state processes that makes thermodynamic analysis difficult. In addition to this, relatively little is usually known about their structure, even in the folded state, since they are less amenable to high resolution crystallographic studies. Consequently, theoretical analysis of their folding interaction is less secure. Amino acid sidechains in such proteins may frequently remain exposed to solvent, on the outside of the elongated chain structure, even in the folded state - so factors such as burying of hydrophobic groups should be of less significance.

Early work on the collagen family of proteins, based on variations in experimental T_m values for a range of naturally-occurring tropocollagens with varying proline and hydroxyproline contents, showed indirectly that backbone hydrogen bonding between polypeptide chains in this triple-stranded structure is unlikely to be the dominant stabilising force (Cooper, 1971, and references therein). These proteins are unusual in containing large numbers of proline and hydroxyproline residues at regular positions in their primary structures, and the number of available inter-chain peptide H-bonding groups will decrease with increasing imino acid content. Paradoxically the estimated heat of unfolding (ΔH_{unf}) increases with increasing pro + hypro content, i.e. unfolding of the collagen triple helix becomes more endothermic the fewer the number of inter-chain hydrogen bonds. The increased thermal stability of collagens with higher pro + hypro content comes mainly from the decrease in rotational degrees of freedom of the unfolded chains, because of the restrictions in backbone rotations imposed by the pyrrolidine ring structure of the proline or hydroxyproline sidechain. This reduces the conformational entropy of the unfolded chain and, indirectly therefore, stabilises the folded structure. The additional enthalpic contributions seem to come from regular solvation effects, possibly involving extended hydrogen-bonded chains of water molecules acting as a sort of "aqueous scaffolding" at the surface of the triple helix. Such interactions are impossible to model or mimic in small molecule systems, and therefore difficult to characterise thermodynamically. More recent calorimetric and other work (reviewed in Privalov, 1982) has confirmed the anomalous enthalpy behaviour of collagen unfolding and the intimate role of water.

Work on other fibrous proteins is less extensive, with the possible exception of the myosin/tropomyosin family of α-helical coiled-coil proteins (Privalov, 1982). Thermal unfolding of these proteins is a highly non-cooperative process, involving several overlapping transitions over an extended range of temperatures. This probably represents the unfolding of various independent or semi-independent domains in these large proteins, and makes thermodynamic analysis difficult.

8. Membrane Proteins

We expect that the factors governing thermodynamic stability of membrane proteins should, in principle, differ significantly from those for water-soluble proteins. In some ways they might be simpler. Unfolding of a protein totally within the non-polar lipid bilayer would involve non of the complications of aqueous solvation or hydrophobic interactions, and would be dominated presumably by breaking of H-bonds and other polar interactions in the folded protein. Unfortunately this neglects the two-phase nature of the system in which membrane proteins frequently have loops of polypeptide exposed to the aqueous phase and where the extent of exposure may well change during folding/unfolding reactions. Experimental data are sparse because of the intrinsic technical difficulties associated with measurements on membrane proteins, and the lack of comprehensive structural data on such systems makes interpretation difficult. Some calorimetric data on unfolding of rhodopsin and bacteriorhodopsin have been obtained (Miljanich et al., 1985; Kahn et al., 1992), including the role of retinal binding and loop regions. Interestingly it appears, in this case at least, that ligand binding and interhelical loops are less significant for protein stability than the side-by-side interactions between helices within the membrane. But the precise nature of these side-by-side interactions has not yet been established.

9. Finale

Why proteins fold is still a bit of a mystery. That is, the opposing thermodynamic forces are so delicately balanced that it is difficult to decide which, if any, are predominant - and indeed the balance may be different in different proteins. Nevertheless, the more we get into this intriguing problem the more we learn about the nature of biomolecular interactions and how they have been fine tuned during evolution to meet biological needs. Chris Anfinsen himself was often pessimistic about the protein folding problem, expressing it this way: that if there are N proteins in the entire world, then by the time we have solved the structure of (N-1) of them perhaps (and only perhaps!) might we accurately predict the structure of the Nth. We still have some way to go.

References

Anfinsen, C.B. (1973). *Principles that govern the folding of protein chains*. Science 181, 223-230.

Anfinsen, C.B. & Scheraga, H.A. (1975). *Experimental and theoretical aspects of protein folding.* Adv. Protein Chem. 29, 205-300.

Antosiewicz, J., McCammon, J.A. & Gilson, M.K. (1994). *Prediction of pH-dependent properties of proteins.* J.Mol.Biol. 238, 415-436.

Baldwin, R.L. (1986). *Temperature dependence of the hydrophobic interaction in protein folding.* Proc.Natl.Acad.Sci.USA, 83, 8069-8072.

Brandts, J.F. (1964). *The thermodynamics of protein denaturation. I. The denaturation of chymotrypsinogen.* J.Amer.Chem.Soc. 86, 4291-4301.

Brandts, J.F. & Lin, L.-N. (1990). *Study of strong to ultratight protein interactions using differential scanning calorimetry*. Biochemistry 29, 6927-6940.

Brooks, C.L., Karplus, M. & Pettitt, B.M. (1988). *Proteins: a theoretical perspective of dynamics, structure, and thermodynamics*. Wiley Interscience, New York.

Chan, H.S., Bromberg, S. & Dill, K.A. (1995). *Models of cooperativity in protein folding.* Phil.Trans.R.Soc. Lond.B 348, 61-70.

Cooper, A. (1971).*Thermal Stability of Tropocollagens - Are Hydrogen Bonds Really Important ?* J.Mol.Biol. 55, 123-127.

Cooper, A. (1976).*Thermodynamic Fluctuations in Protein Molecules*. Proc.Natl.Acad.Sci.USA, 73, 2740-2741.

Cooper, A. (1984).*Protein Fluctuations and the Thermodynamic Uncertainty Principle*. Prog.Biophys.Molec.Biol. 44, 181-214.

Cooper, A. (1992). *Effect of cyclodextrins on the thermal stability of globular proteins.* J.Amer.Chem.Soc. 114, 9208-9209.

Cooper, A., Eyles, S.J., Radford, S.E. & Dobson, C.M. (1992). *Thermodynamic consequences of the removal of a disulphide bridge from hen lysozyme*. J.Mol.Biol 225, 939-943.

Cooper, A. & McAuley-Hecht, K.E. (1993). *Microcalorimetry and the Molecular Recognition of Peptides and Proteins.* Phil.Trans.R.Soc. Lond. A 345, 23-35.

Creighton, T.E. (1988). *Disulphide bonds and protein stability.* BioEssays 8, 57-63.

Creighton, T.E. (1991). *Stability of folded conformations.* Curr.Opin.Struct.Biol. 1, 5-16.

Creighton, T.E. (1992) Ed.: *Protein Folding* (W.H. Freeman & Co., New York).

Dao-pin, S., Anderson, D.E., Baase, W.A., Dahlquist, F.W. & Matthews, B.W. (1991). *Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme.* Biochemistry 30, 11521-11529.

Dill, K.A. (1990a). *Dominant forces in protein folding.* Biochemistry 29, 7133-7155.

Dill, K.A. (1990b). *The meaning of hydrophobicity.* Science 250, 297.

Dill, K.A. (1993). *Folding proteins: finding a needle in a haystack.* Curr.Opin.Struct. Biol. 3, 99-103.

Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D. and Chan, H.S. (1995). *Principles of protein folding - A perspective from simple exact models.* Protein Sci. 4, 561-602.

Dill, K.A. and Stigter, D. (1995). *Modeling protein stability as heteropolymer collapse.* Adv.Protein Chem. 46, 59-104.

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

Doig, A.J. and Williams, D.H. (1991). *Is the hydrophobic effect stabilizing or destabilizing in proteins ? The contribution of disulphide bonds to protein stability.* J.Mol.Biol. 217, 389-398.

Dunitz, J.D. (1995). *Win some, lose some: Enthalpy-entropy compensation in weak intermolecular interactions.* Chemistry & Biology 2, 709-712.

Edsall, J.T. (1995). *Hsien Wu and the first theory of protein denaturation (1931).* Adv.Protein Chem. 46, 1-5.

Eisenberg, D. and Kauzmann, W. (1969). *The Structure and Properties of Water.* (Oxford University Press, 1969).

Fersht, A.R., Matouschek, A. & Serrano, L. (1992). *The folding of an enzyme.1. Theory of protein engineering analysis of stability and pathway of protein folding.* J.Mol.Biol. 224, 771-782.

Flory, P.J. (1956). *Theory of elastic mechanisms in fibrous proteins.* J.Am.Chem.Soc. 78, 5222-5235.

Franks, F. (1995). *Protein destabilization at low temperatures.* Adv.Protein Chem. 46, 105-139.

Fukada, H., Sturtevant, J.M. & Quiocho, F.A. (1983) *Thermodynamics of the binding of L-arabinose and of D-galactose to the L-arabinose-binding protein of Escherichia coli. J.Biol.Chem.* 258, 13193-13198.

Galisteo, M.L., Mateo, P.L. & Sanchez-Ruiz, J.M. (1991). *Kinetic study on the irreversible thermal denaturation of yeast phosphoglycerate kinase.* Biochemistry, 30, 2061-2066.

Gilson, M.K. & Honig, B. (1986). *The dielectric constant of a folded protein.* Bioplymers, 25, 2097-2119.

Griko, Y.V., Freire, E. & Privalov, P.L. (1994). *Energetics of the α-lactalbumin states: a calorimetric and statistical thermodynamic study.* Biochemistry 33, 1889-1899.

Griko, Y.V. & Privalov, P.L. (1994). *Thermodynamic puzzle of apomyoglobin unfolding.* J.Mol.Biol. 235, 1318-1325.

Grunwald, E. and Steel, C (1995). *Solvent reorganization and thermodynamic enthalpy-entropy compensation.* J.Am.Chem.Soc. 117, 5687-5692.

Honig, B., Sharp, K.A. & Yang, A.-S. (1993). *Macroscopic models of aqueous solutions: biological and chemical applications.* J.Phys.Chem. 97, 1101-1109.

Honig, B. & Yang, A.-S. (1995). *Free energy balance in protein folding.* Adv.Protein Chem. 46, 27-58.

Hu, C-Q., Sturtevant, J.M., Thomson, J.A., Erickson, R.E. & Pace, C.N. (1992). *Thermodynamics of ribonuclease T1 denaturation.* Biochemistry, 31, 4876-4882.

Jackson, W.M. & Brandts, J.F. (1970). *Thermodynamics of protein denaturation. Calorimetric study of the reversible denaturation of chymotrypsinogen and conclusions regarding the accuracy of the two-state approximation.* Biochemistry, 9, 2294-2301.

Jacobson, H. & Stockmayer, W.H. (1950). *Intramolecular reaction in polycondensations. I. The theory of linear systems.* J.Chem.Phys. 18, 1600-1606.

Johnson, C.M., Cooper, A. & Stockley, P.G. (1992). *Differential scanning calorimetry of thermal unfolding of the methionine repressor protein (MetJ) from Escherichia coli.* Biochemistry 31, 9717-9724.

Johnson, C.M. & Fersht, A.R. (1995). *Protein stability as a function of denaturant concentration: the thermal stability of barnase in the presence of urea.* Biochemistry 34, 6795-6804.

Johnson, R.E., Adams, P., and Rupley, J.A. (1978). *Thermodynamics of protein cross-links.* Biochemistry, 17, 1479-1484.

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

Kahn, T.W., Sturtevant, J.M. & Engelman, D.M. (1992). *Thermodynamic measurements of the contributions of helix-connecting loops and of retinal to the stability of bacteriorhodopsin.* Biochemistry 31, 8829-8839.

Khechinashvili, N.N., Janin, J. & Rodier, F. (1995). *Thermodynamics of the temperature-induced unfolding of globular proteins.* Protein Sci. 4, 1315-1324.

Klotz, I.M. & Farnham, S.B. (1968). *Stability of an amide-hydrogen bond in an apolar environment.* Biochemistry, 7, 3879-3882.

Kresheck, G.C. & Klotz, I.M. (1969). *The thermodynamics of transfer of amides from an apolar to an aqueous solution.* Biochemistry, 8, 8-12.

Kuroki, R., Inaka, K., Taniyama, Y., Kidokoro, S., Matsushima, M., Kikuchi, M., and Yutani, K. (1992). *Enthalpic destabilization of a mutant human lysozyme lacking a disulfide bridge between cysteine-77 and cysteine-95.* Biochemistry, 31, 8323-8328.

Lazaridis, T., Archontis, G. & Karplus, M. (1995). *Enthalpic contributions to protein stability: insights from atom-based calculations and statistical mechanics.* Adv. Protein Chem. 47, 231-306.

Lee, B. (1991). *Isoenthalpic and isoentropic temperatures and the thermodynamics of protein denaturation.* Proc.Natl.Acad.Sci.USA, 88, 5154-5158.

Lepock, J.R., Ritchie, K.P., Kolios, M.C., Rodahl, A.M., Heinz, K.A., and Kruuv, J. (1992). *Influence of transition rates and scan rate on kinetic simulations of differential scanning calorimetry profiles of reversible and irreversible protein denaturation.* Biochemistry 31, 12706-12712.

Levinthal, C. (1968). *Are there pathways for protein folding ?* J.Chim.Phys. 65, 44-45.

Makhatadze, G.I. & Privalov, P.L. (1992). *Protein interactions with urea and guanidinium chloride: a calorimetric study.* J.Mol.Biol. 226, 491-505.

Makhatadze, G.I. & Privalov, P.L. (1993). *Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration.* J.Mol.Biol. 232, 639-659.

Makhatadze, G.I. & Privalov, P.L. (1995). *Energetics of protein structure.* Adv. Protein Chem., 47, 307-425.

Makhatadze, G.I. & Privalov, P.L. (1996). *On the entropy of protein folding.* Protein Sci. 5, 507-510.

Matouschek, A., Matthews, J.M., Johnson, C.M. & Fersht, A.R. (1994). *Extrapolation to water of kinetic and equilibrium data for the unfolding of barnase in urea solutions.* Protein Eng., 7, 1089-1095.

Miljanich, G., Brown, M., Mabrey-Gaud, S., Dratz, E. & Sturtevant, J.M. (1985). *Thermotropic behavior of retinal rod membranes and dispersions of extracted phospholipids.* J.Membr.Biol. 85, 79-86.

Mitraki, A., Betton, J.M., Desmadril, M, and Yon, J.M. (1987). *Quasi-irreversibility in the unfolding-refolding transition of phosphoglycerate kinase induced by guanidine hydrochloride.* Eur.J.Biochem. 163, 29-34.

Murphy, K.P. & Freire, E. (1992). *Thermodynamics of structural stability and cooperative folding behavior in proteins.* Adv.Protein Chem. 43, 313-361.

Murphy, K.P., Privalov, P.L. & Gill, S.J. (1990). *Common features of protein unfolding and dissolution of hydrophobic compounds.* Science 247, 559-561.

Naghibi, H., Tamura, A, and Sturtevant, J.M. (1995). *Significant discrepancies between van't Hoff and calorimetric enthalpies.* Proc.Natl.Acad.Sci. USA, 92, 5597-5599.

Pace, C.N., Grimsley, G.R., Thomson, J.A., and Barnett, B.J. (1988). *Conformational stability and activity of ribonulease T1 with zero, one, and two intact disulfide bonds.* J.Biol.Chem. 263, 11820 -11825.

Pfeil, W. & Privalov, P.L. (1976a). *Thermodynamic investigations of proteins. I. Standard functions for proteins with lysozyme as an example.* Biophys.Chem. 4, 23-32.

Pfeil, W. & Privalov, P.L. (1976b). *Thermodynamic investigations of proteins. II. Calorimetric study of lysozyme denaturation by guanidine hydrochloride.* Biophys.Chem. 4, 33-40.

Pfeil, W. & Privalov, P.L. (1976c). *Thermodynamic investigations of proteins. III. Thermodynamic description of lysozyme.* Biophys.Chem. 4, 41-50.

Poland, D.C. and Scheraga, H.A. (1965). *Statistical mechanics of noncovalent bonds in polyamino acids. VIII. Covalent loops in proteins.* Biopolymers 3, 379-399.

Privalov, P.L. (1979). *Stability of proteins: Small globular proteins.* Adv.Protein Chem. 33, 167-241.

Privalov, P.L. (1982). *Stability of proteins: Proteins which do not present a single cooperative system.* Adv.Protein Chem. 35, 1-104.

Privalov, P.L. (1990). *Cold denaturation of proteins.* Crit.Rev.Biochem.Mol.Biol. 25, 281-305.

Privalov, P.L. (1992). *Physical basis of the stability of the folded conformations of proteins.* In: Creighton (1992).

Privalov, P.L. (1996). *Intermediate states in protein folding.* J.Mol.Biol. 258, 707-725.

Privalov, P.L. & Gill, S.J. (1988). *Stability of protein structure and hydrophobic interactions.* Adv.Protein Chem. 39, 191-234.

Privalov, P.L., Gill, S.J. & Murphy, K.P. (1990). *The meaning of hydrophobicity (response).* Science 250, 297-298.

Privalov, P.L. & Khechinashvili, N.N. (1974). *A thermodynamic approach to te problem of stabilization of globular protein structure: a calorimetric study.* J.Mol.Biol. 86, 665-684.

Privalov, P.L. & Makhatadze, G.I. (1990). *Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polypeptide chain of proteins.* J.Mol.Biol. 213, 385-391.

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

Privalov, P.L. & Potekhin, S.A. (1986) *Scanning calorimetry in studying temperature-induced changes in proteins.* Methods Enzymol. 131, 4-51.

Ptitsyn, O.B. (1995). *Molten globule and protein folding.* Adv.Protein Chem. 47, 83-229.

Radford, S.E., Woolfson, D.N., Martin, S.R., Lowe, G. & Dobson, C.M. (1991). *A three-disulphide derivative of hen lysozyme: structure, dynamics and stability.* Biochem.J. 273, 211-217.

Rose, G.D. & Wolfenden, R. (1993). *Hydrogen-bonding, hydrophobicity, packing and protein folding.* Annu.Rev.Biophys.Biomol.Struct. 22, 381-415.

Sanchez-Ruiz, J.M., Lopez-Lacomba, J.L., Cortijo, M. & Mateo, P.L. (1988). *Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin.* Biochemistry, 27, 1648-1652.

Santoro, M.M. & Bolen, D.W. (1992). *A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range.* Biochemistry 31, 4901-4907.

Santoro, M.M., Liu, Y., Khan, S.M.A., Hou, L-X. & Bolen, D.W. (1992). *Increased thermal stability of proteins in the presence of naturally occuring osmolytes.* Biochemistry 31, 5278-5283.

Schellman, J.A. (1955). *The stability of hydrogen-bonded peptide structures in aqueous solution.* C.R.Trav.Lab.Carlsberg Ser.Chim. 29, 230-259.

Schellman, J.A. (1987a). *The thermodynamic stability of proteins.* Annu.Rev.Biophys.Chem. 16, 115-137.

Schellman, J.A. (1987b). *Selective binding and solvent denaturation.* Biopolymers 26, 549-559.

Schmid, F.X., Mayr, L.M., Mücke, M. & Schönbrunner, E.R. (1993). *Prolyl isomerases: role in protein folding.* Adv.Protein Chem. 44, 25-66.

Serrano, L., Kellis, J.T., Cann, P., Matouschek, A. & Fersht, A.R. (1992). *The folding of an enzyme.2. Substructure of barnase and the contribution of different interactions to protein stability.* J.Mol.Biol. 224, 783-804.

Spolar, R.S., Livingstone, J.R. & Record, M.T. (1992). *Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water.* Biochemistry 31, 3947-3955.

Stein, R.L. (1993). *Mechanism of enzymatic and nonenzymatic prolyl cis-trans isomerization.* Adv.Protein Chem. 44, 1-24.

Stigter, D. & Dill, K.A. (1990). *Charge effects on folded and unfolded proteins.* Biochemistry, 29, 1262-1271.

Sturtevant, J.M. (1974). *Some applications of calorimetry in biochemistry and biology.* Ann.Rev.Biophys.Bioeng. 3, 35-51.

Sturtevant, J.M. (1987). *Biochemical applications of differential scanning calorimetry.* Ann.Rev.Phys.Chem. 38, 463-488.

Tanford, C. (1968). *Protein denaturation.* Adv. Protein Chem. 23, 121-275.

Tanford, C. (1970). *Protein denaturation.* Adv. Protein Chem. 24, 1-95.

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

Thornton, J.M. (1981). *Disulfide bridges in globular proteins.* J.Mol.Biol. 151, 261-287.

Timasheff, S.N. (1992). *Water as ligand: preferential binding and exclusion of denaturants in protein unfolding.* Biochemistry 31, 9858-9864.

Velicelebi, G. & Sturtevant, J.M. (1979). *Thermodynamics of the denaturation of lysozyme in alcohol-water mixtures.* Biochemistry 18, 1180-1186.

Vogl, T., Brengelmann, R., Hinz, H-J., Scharf, M., Lötzbeyer, M. & Engels, J.W. (1995). *Mechanism of protein stabilization by disulfide bridges: calorimetric unfolding studies on disulfide-deficient mutants of the α-amylase inhibitor Tendamistat.* J.Mol.Biol. 254, 481-496.

Weber, G. (1975). *Energetics of ligand binding to proteins.* Adv.Protein Chem. 29, 1-83.

Weber, G. (1993). *Thermodynamics of the association and the pressure dissociation of oligomeric proteins.* J.Phys.Chem., 97, 7108-7115.

Weber, G. (1995). *van't Hoff revisited: Enthalpy of association of protein subunits.* J.Phys.Chem., 99, 1052-1059.

Woolfson, D.N., Cooper, A., Harding, M.M., Williams, D.H. & Evans, P.A. (1993) *Protein folding in the absence of the solvent ordering contribution to the hydrophobic interaction.* J.Mol.Biol. 229, 502-511.

Wu, H. (1931). *Studies on denaturation of proteins XIII. A theory of denaturation.* Chinese Journal of Physiology, 5, 321-344. (Reprinted in Adv.Protein Chem. (1995) 46, 6-26.)

Wyman, J. (1964). *Linked functions and reciprocal effects in hemoglobin: a second look.* Adv.Protein Chem. 19, 223-286.

Wyman, J. & Gill, S.J. (1990). *Binding and Linkage: functional chemistry of biological macromolecules.* (University Science Books, Mid Valley, CA, 1990).

Yancey, P.H., Clark, M.E., Hand, S.C., Bowlus, R.D. & Somero, G.N. (1982). *Living with water-stress: evolution of osmolyte systems.* Science 217, 1214-1222.

Yang, A.-S., Gunner, M.R., Sampogna, R., Sharp, K. & Honig, B. (1993). *On the calculation of pK_as in proteins.* Proteins:Structure,Function and Genetics, 15, 252-265.