THERMOSTABILITY OF ENZYMES

Thermostability of Enzymes

EDITOR
M.N. Gupta

Springer-Verlag

Narosa Publishing House

EDITOR
M.N. Gupta
Chemistry Department
Indian Institute of Technology, Delhi
New Delhi, India

ISBN 3-540-56275-3 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-56275-3 Springer-Verlag New York Berlin Heidelberg
ISBN 81-85198-49-7 Narosa Publishing House New Delhi

Printed in India at Rajkamal Electric Press, Delhi 110033.

Thermostability of Enzymes

Springer-Verlag
Berlin
Heidelberg
New York
London
Paris
Tokyo
Hong Kong
Barcelona
Budapest

Narosa Publishing House
New Delhi
Madras
Bombay
Calcutta

Preface

Enzyme stability has always engaged the attention of workers from different areas: enzymologists have naturally always preferred stable enzymes; physical chemists who did much of the early work on denaturation and microbiologists who found that all enzymes were not created equal when it came to stability. In recent years, biochemical engineers, molecular biologists and biotechnologists have joined this evergrowing list. Also, as enzymes find increasing number of applications in both synthesis and analysis, a large number of industries are now concerned with both producing and utilizing enzymes in the stable form. An important example in this regard is PCR technology which would not have been practical but for the availability of a thermostable DNA polymerase. (The technique is not discussed in this book as it is comprehensively dealt with at many other places e.g. PCR: A Practical Approach M.J. McPherson, P. Quirke and G.R. Taylor, Oxford University Press, Oxford, 1991). Thus, today enzyme stability is an important issue which straddles both basic and applied research. It is now a truely multidisciplinary area. Of all the harsh conditions, stability towards high temperature is perhaps the most important one. Also, a thermostable enzyme is often likely to be stable in general towards other denaturing conditions.

This book is an effort to put together at a single place, the various facets of enzyme thermostability and thermostabilization. It is felt that not only it would provide a ready reference for people working in different areas (but concerned with enzyme thermostabiiity), it would also act as a catalyst for the further growth of the subject.

When a multiauthored book (such as this one), has an extremely sharp focus on a single subject, it is difficult to avoid some overlap and repetition. I have intentionally let this happen occasionally in this book, partly so as not to disturb the continuity in the discussion and partly because some important concepts are worth repeating especially in different contexts.

The first seven chapters predominantly deal with understanding thermostability while the last four chapters describe different techniques of thermostabilization. In the last chapter, the authors have rightly emphasized that these twin areas—understanding thermostability and thermostabilization—are highly interdependent. Thus, it is unavoidable that first chapter on protein structure and stability should refer to results obtained with protein engineering!

I am grateful to the contributing authors who joined me in this endeavour. This is a joint effort. However, if there are mistakes and deficiencies, I take the blame.

It is also a pleasure to acknowledge many others who helped. The encouragement by my friend Prof. Bo Mattiasson (who also readily agreed to co-author Chapter 2 which tells us how we may obtain reasonable amount of an enzyme for practical

applications before we proceed to stabilize it) has been very important to me in my task. Both, Council of Scientific and Industrial Research, India and Department of Science and Technology (Government of India), India have funded my research work over the years. Without their support, it would not have been possible for me to sustain fascination for thermostability. I thank IIT Delhi authorities for their book writing grant.

I wish to acknowledge the help rendered by Prof. J.C. Ahluwalia by reviewing Chapter 6.

I also wish to thank members of my research group, both past as well as present, who by their work, fuelled my interest in this area. Extra thanks are due to Dr. Sarfraz Ahmad, Dr. Shuchita Kumar, Dr. Renu Tyagi, Dr. Swarita Gopal, Ms. Renu Batra, Ms. Ritu Aggarwal, Ms. Nidhi Aneja and Ms. Namrata Pathak for helping me in many ways during editing. I was helped in the index preparation by Ms. Nidhi Aneja and Ms. Namrata Pathak.

I thank my wife, Sulbha, for her critical support. My son, Chetan, about 9 years old, encouraged me by showing constant interest in this book and occasionally grilling me by asking questions like what are enzymes and from where do they come!

Finally, I am thankful to people at Narosa Publishing House, particularly Mr N.K. Mehra and Mr M.S. Sejwal, who showed lot of patience and gave necessary advice whenever needed.

I would welcome suggestions and comments on this book.

Indian Institute of Technology, Delhi
New Delhi, India.

M.N. GUPTA

Contents

...the structures were austere and simple, until one looked at them and realized what work, what complexity of method, what tension of thought had achieved the simplicity.....

...structure is a solved problem of tension, of balance, of security in counter thrusts.

—Ayn Rand

1. Protein Structure and Stability

R. Sowdhamini and P. Balaram
Molecular Biophysics Unit
Indian Institute of Science
Bangalore, India

The enormous structural diversity of proteins coupled with their functional versatility has made the field of protein structure and folding a central area of modern structural biology [1]. The increasing use of proteins in biotechnology has provided a fresh impetus to studies relating their structure and stability [2]. The precise relationships between protein sequence and three-dimensional structure remain to be defined. A universal 'folding code', which will permit definitive structural predictions from sequence data, is the Holy Grail of contemporary molecular biophysics [3–5]. Advances in crystallography, NMR spectroscopy, computational methods and molecular biology have provided a wide range of techniques to attack the problem of defining relationships between protein structure, folding patterns, stability and biological function.

The native folded structures of globular proteins are only marginally stabilised with respect to their unfolded forms ($\Delta\Delta G$ unfolding ~ 5–20 kcalmol^{-1}) [6, 7]. While hydrophobic effects [8] contribute predominantly to the acquisition of compact folded structures, by polypeptide chains in aqueous solution, the loss of conformational entropy tends to destabilise the folded state [9, 10]. The native structure with its detailed architecture, which permits maximisation of favourable interactions, is a result of a delicate balance between these opposing forces [10]. Recent studies of proteins, enzymes in particular, employing the techniques of site directed mutagenesis have revealed that an enormous number of sequence changes are indeed structurally and functionally tolerated, leading to the concept of plasticity of protein structures [11, 12]. The relevant question in studies of protein folding today, is not—What three-dimensional structure will a given polypeptide sequence adopt?—rather it is—What are the various sequences that are compatible with a given folding pattern [13]? A particularly attractive approach to assess the importance of specific molecular interactions in determining the thermal stability of folded conformations is to examine the effects of carefully chosen mutations on the denaturation temperatures (t_m) of structurally well characterised proteins [14, 15]. Research in this area is also fuelled by the need to rationally design functionally viable mutant enzymes, with high thermal stabilities for practical applications.

The thermal stability of enzymes can have important consequences for the function and survival of complex organisms. Thermophilic microorganisms are

indeed capable of existence at temperatures as high as 100°C, a property that places special requirements on the structural stability of their proteins [16]. Quite often disease conditions in humans are associated with single point mutations in key enzymes which lead to thermolability. Representative examples include single residue alterations in the enzymes triosephosphate isomerase [17], glucose-6-phosphate dehydrogenase [18] or aldolase [19] which lead to hemolytic anemia. While the physiological consequences of naturally occurring mutations have proved a valuable guide to deciphering structure-stability relationships, recent purposeful attempts at engineering protein structure promise to provide valuable new insights in this area.

PROTEIN UNFOLDING AND DENATURATION

The term 'denaturation' was first introduced into the literature by Chick and Martin who indeed distinguished this process from the phenomenon of heat induced aggregation [20]. However, the view of denaturation as an unfolding process is attributed to Wu, who in a remarkably prescient manner recognised that this event must involve structural disorganisation, at a time when definitive ideas regarding protein structure were yet to be formulated [21]. In the six decades since, denaturation has come to be regarded as the process of transformation of a protein to a non-native, and in the case of enzymes to an inactive conformation. X-ray crystallography has successfully permitted visualisation of the unique globular structures of a large number of proteins. These studies coupled with the demonstration of enzymatic activity in crystals has led to the current view that crystalline conformations present an accurate representation of the native structure of proteins [22]. For a large number of examples, the thermodynamic reversibility of the denaturation process has been established, suggesting that quite often the native folded form is indeed a global energy minimum [10]. Recent studies on assisted folding processes, involving 'molecular chaperones' appear to suggest that folding can sometimes be externally catalysed, a feature which may be of importance for *in vivo* refolding [23, 24]. However, the overwhelming majority of protein folding studies are confined to examples where *in vitro* refolding is achievable. The extremely small energy difference between folded and unfolded forms of proteins suggests that protein structures are finely poised, with a delicate balance of intramolecular and solvation interactions favouring the native conformation [6, 7]. The multiplicity of interactions—hydrogen bonding [25], electrostatic forces [26], hydrophobic effects [8, 27], packing contributions [28] and solvation [29]—that stabilise protein structures renders studies of structure-stability relationships very complex. Recent attempts to dissect the contributions of specific interactions to protein folding and stability have focussed on model proteins, which have the advantage of being structurally well characterised by high resolution X-ray structures and have been cloned into high expression systems that permit preparation of mutant proteins. Representative examples include bacteriophage T4-lysozyme [30], subtilisin [31], barnase [32], λ-repressor [33], staphylococcal nuclease [34] and gene V protein of bacteriophage f1 [35]. A complicating factor in most studies of protein unfolding is the absence of any structural characterisation of the

unfolded state, which often maintains a considerable degree of residual structure. Denaturation studies generally involve either variation of temperature (heat or cold denaturation) or addition of chaotropic agents as a means of inducing unfolding [36, 37]. The non-native structures so generated can be appreciably heterogeneous. Thus, while mutational effects on native structures can be modelled with some degree of confidence, the effects of structural changes on unfolded forms are much harder to assess. Kauzmann suggests that a random coil can adopt Z^x conformations, where Z = number of degenerate rotational states and x = number of rotatable bonds in a polypeptide [38]. However, excluded volume effects alone would largely reduce the number of possible conformations in an unfolded state [9, 39]. The unfolded state of a protein contributes appreciably to the overall entropy of a protein [40]. Shortle and Meeker have put forward a thermodynamic test for assigning selective effects of amino acid substitution to the unfolded state [41]. At normal temperatures, ΔC_p for the unfolding of a protein is large [42] and positive (~ 1–2 kcalmol^{-1}), which confirms the central role of hydrophobic effects in protein stabilisation [43]. Denaturation in a protein can be induced by cold as well and was first observed in β-lactoglobulin [44]. This is unsurprising in view of the competing effect of hydrophobic interactions and chain entropy in determining the stability of the folded state [9].

Once proteins are thermally unfolded, many factors might lead to irreversible damage to the protein [45–51]. Genetic screens for selecting thermally stable mutations prove to be useful in correlating *in vivo* stability with *in vitro* experiments [52]. A structure-sequence comparison of homologous proteins like hemoglobin and cytochromes from different organisms reveal that invariant sites are present in the interior of a protein [53–55]. Comparisons of thermophilic and mesophilic proteins, which differ in one or two residues, provide better insights into the problem of protein stability [56–58]. The remarkably little difference in ΔG of the folded form from an unfolded one implies that even a change in one hydrogen bond or a methylene group, at the appropriate position, will be enough to offset the balance. However, since the mechanisms governing these changes remain obscure, empirical approaches are resorted to improve and also to understand protein stability.

Table 1 lists several approaches which have been examined with the goal of enhancing thermal stability of native protein structures. The most promising strategies using rational mutagenesis involve alterations of interior packing, modification of electrostatic interactions or the introduction of covalent cross-links particularly disulfide bridges. A few illustrative examples of attempts to probe structure-stability relationships are discussed below.

PACKING EFFECTS

The fact that soluble proteins can achieve globular structures necessarily implies that protein interiors are well packed with solvophobic residues. The high packing density of protein cores results in relatively few internal water molecules [10]. Although, optimum filling of available space is a principal requirement for interior core stability, the role of subtle electrostatic effects as in the case of quadrupole-quadrupole interactions in aromatic sidechain packing has also been

Table 1. Representative examples of procedures to enhance thermal stability of proteins

Strategy	Reference	Protein	Remarks
1. (a) Introduction of disulfide bonds	Matsumura et al (1989) [59]	T4 lysozyme	Three disulfide bonds introduced at flexible regions of the protein
	Sauer et al (1986) [60]	λ-Repressor	A disulfide bond was introduced at the dimer interface
(b) Chemical modification at reactive sidechains	Gupta (1991) [61]	—	General review
	Shiksnis et al (1986) [62]	α-Chymotrypsin	Attachment of pyromellitic dianhydride to free amino groups leads to high thermal stability
2. Alteration of packing arrangements in the hydrophobic core and elimination of interior cavities	Yutani et al (1987) [63]	α-Subunit of Trp synthase	Gln 49 found in the interior of the protein was replaced by other amino acids
	Kellis et al (1988) [64]	Barnase	Ile 96 in the hydrophobic core was replaced by Val or Ala
	Matsumura et al (1988a) [65]	Kanamycin nucleotidyl transferase	Buried Asp 80 was replaced by Tyr
	Matsumura et al (1988b) [66]	T4 lysozyme	Ile 3 was substituted by other amino acids
	Lim and Sauer (1989) [67]	λ-Repressor	Casette mutagenesis technique was adopted to alter sets of residues, randomly in the N-terminal domain
	Lim and Sauer (1991) [68]	λ-Repressor	Val 36, Met 40 and Val 147 were replaced by random mutagenesis
	Shortle et al (1990) [69]	Staphylococcal nuclease	All the Leu, Val, Tyr, Phe, Met and Ile residues were mutated one at a time to Gly/Ala

(Contd.)

Strategy	Reference	Protein	Remarks
	Daopin et al (1991) [70]	T4-lysozyme	Residues at helix interface were replaced by others
	Sandberg and Terwilliger (1989) [71]	Gene V Protein	Val 35 and Ile 47, in the interior of the protein and in mutual contact, were changed to Ile and Val, respectively
3. Engineering of helix-dipole interactions	Sali et al (1988) [72]	Barnase	Histidine was introduced at position 18, in order to stabilise the C-cap region of a helix
	Seranno and Fersht (1989) [73]	Barnase	Thr residues were introduced at positions 6 and 26, at the N-cap regions of two helices
	Nicholson et al (1988) [74]	T4-lysozyme	Two Asp residues introduced near helix terminii
4. Introduction of new hydrogen bonds and salt bridge interactions	Alber et al (1987) [75]	T4-lysozyme	Thr 157 was replaced by other amino acids
	Perry et al (1987) [76]	Dihydrofolate reductase	Thr 113 → Val led to a decrease in stability
	Erwin et al (1987) [77]	Subtilisin BPN′	Salt-bridges were engineered at specific sites
5. Reduction of backbone conformational entropy	Matthews et al (1987) [78]	T4-lysozyme	Gly 77 in an α-helix was replaced by Ala → Pro
	Hecht et al (1986) [79]	λ-Repressor	Two Gly residues in helices were replaced
	Imanaka et al (1986) [80]	Neutral protease of *B. stearothermophilus*	Gly 61 and Gly 144 were replaced by Ala
6. Mutations based on analogy with homologous proteins from thermophilic organisms	Imanaka et al (1986) [80]	Neutral protease	Replacements were based on thermolysin as a model protein
	Takagi et al (1990) [81]	Subtilisin E	A disulfide bond was introduced based on a thermophilic homologue, aqualysin

7. Modification of surface residues	Pakula and Sauer (1990) [82]	λ-Cro	Hydrophobic Tyr 26 on the surface was replaced by Cys
	Ahern et al (1987) [83]	Triose phosphate isomerase	Surface Asn residues, 14 and 28 were replaced
	Perry and Wetzel (1987) [84]	T4-lysozyme	Surface Cys residues, 54 and 97 were replaced by Val and Thr respectively
	Estel et al (1985) (1985) [85]	Subtilisin	Met 222 was replaced by Gln
	Hecht et al (1984) [86]	λ-Repressor	Lys 4 and Gln 33 were replaced by Gln and Tyr respectively
8. Immobilisation	Klibanov (1979, 1983) [87, 88]	—	General reviews

invoked [26]. The driving force for the condensation of the polypeptide chain into a globular fold is hydrophobic (solvophobic in a more general sense), as recognised by Kauzmann in the 1950s [27]. The knowledge of the structural details of residue packing in protein cores has stimulated several attempts to evaluate the precise role of sidechain packing interactions in determining stability of protein structures [89]. In using mutagenesis to probe packing effects, the effects of hydrophobicity changes must be distinguished from genuine packing changes. A study of the gene V protein from bacteriophage f1 focussed on two internal residues that make close contact, Val 35 and Ile 47 [71]. Mutants involving Val → Ile and Ile → Val at both positions were examined with respect to guanidine hydrochloride denaturation. All the mutant proteins were less stable compared to the wild type but no clear conclusions regarding the effects of substitution at buried sites could be reached [71]. Lim and Sauer have used cassette mutagenesis to alter randomly, residues within the hydrophobic core of the N-terminal domain of λ-repressor [67]. Using a genetic screen functional mutants could be isolated. This analysis confirms that "hydrophobicity is the most essential feature of the core sequence", although not all hydrophobic core sequence fold correctly [67]. A further analysis of the stability of 78 mutant proteins, generated by a random mutagenesis strategy (residues Val 36, Met 40, Val 47 of λ-repressor were randomised using codons for Val, Leu, Ile, Met and Phe), showed that about 70% of the isolated sequences have some biological activity [68]. This suggests that a large fraction of the mutants do indeed maintain the basic structure of the repressor. The wild type (WT) protein has a melting temperature (t_m) of 56°C, while the mutants have t_m values ranging from 47 to 60°C. This corresponds to differences in the ΔG of unfolding ranging from –1.3 kcalmol^{-1} to 0.5 kcalmol^{-1}, relative to WT protein. Interestingly, only two mutants (V36I and V36L/M40L/V47I) have t_m slightly higher than WT protein [68]. A detailed examination of the contribution of hydrophobic residues to stability in the case of staphylococcal nuclease focussed on 83 mutants obtained by mutagenesis of 11 Leu, 9 Val, 7 Tyr, 5 Ile, 4 Met and 3 Phe residues singly to Ala and Gly residues [69]. Additional Ile → Val mutants were also studied. Again, no mutants that showed enhanced stability were isolated. The major conclusion of this analysis is that residue substitutions may "destabilise a protein indirectly via their effects on the structure and free energy of the denatured state". Shortle et al, indeed conclude rather provocatively that "packing of hydrophobic residues in the native state should be considered more a vestige of dominant interactions that occurred in the denatured state than an ongoing source of stability that acts exclusively of the native state" [69].

A feature of many analyses of packing effects on protein structures is the absence of detailed structural studies on the consequences of specific mutations. A recent attempt at correlating structure and thermodynamic data addresses the problem of packing of two internal helices in T4 lysozyme [70]. Ala 98 on helix 94–107 nestles up closely to residues Val 149 and Thr 152 on helix 143–153. Figure 1 shows a view of the two packed helices with the relevant sidechains indicated. The A98V mutation had already been identified as a temperature sensitive lesion, in a random genetic screen, the denaturation temperature being

lowered by 15°C ($\Delta G = -4.9$ kcalmol^{-1}). The crystal structure of this mutant reveals that accommodation of the bulkier Val residue requires some atoms on the two helices to move apart by 0.7 Å. Attempts to compensate for the A98V mutations were made by mutating residues 149 and 152 (Thr 152 → Ser, Val 149 → Cys/Ile). High resolution crystal structures of double and multiple mutants were determined, as were thermal denaturation profiles. The results of this investigation suggest that structure-stability relationships will prove hard to establish unambiguously [64]. Protein structures can adjust in subtle fashion over considerable distances from the site of a mutation. Consequently, the structural context of a mutation will determine the effect on protein stability. Functionally, similar proteins can then have hydrophobic core structures that can be quite different in detail. Packing calculations of the type developed by Ponder and Richards [89] suffer from an overly restrictive set of packing conditions, since structural relaxations are not easy to predict [70]. It would appear from these studies of hydrophobic core mutations that rational engineering of protein interiors to enhance stability may prove a more difficult task than originally anticipated.

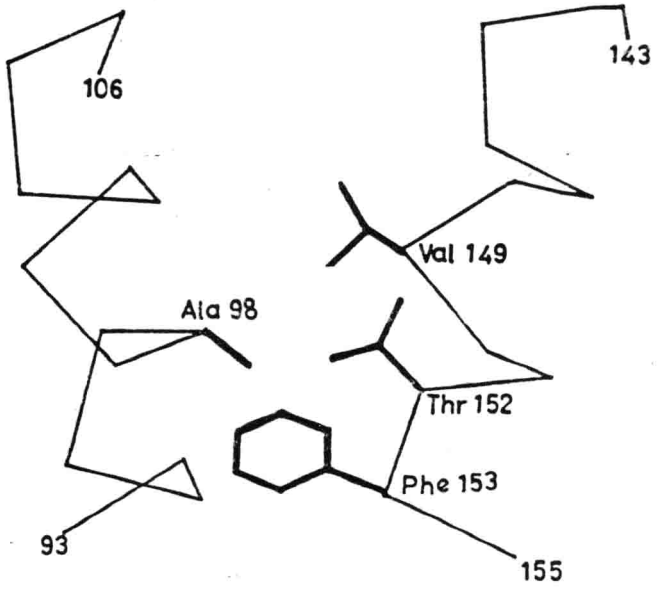

Fig. 1 Helix interface of the 93-106 and 143-155 helices of T4 lysozyme. Only the C^{α} atoms are shown as representative of each residue. The sidechain atoms of Ala 98, Val 149, Thr 152 and Phe 153 are also shown. The four residues were found to be in the interior of the protein and closely packed. Amino acid substitutions in this region lead to thermal destabilisation of the protein and subtle structural changes [70].

ELECTROSTATIC INTERACTIONS AND HYDROGEN BONDS

While hydrophobic effects dominate the folding of proteins into native globular structures, electrostatic interactions and hydrogen bonds contribute to overall stabilisation. The number of ion pairs observed in protein structures is relatively