THERMOSTABILITY AND THERMOACTIVITY OF EXTREMOZYMES

Michael J. Danson and **David W. Hough**
University of Bath, UK

Keywords: archaea, catalysis, citrate synthase, compatible solutes, conformational flexibility and rigidity, directed evolution, enzymes, extremophiles, extremozymes, glutamate dehydrogenase, hydrophobic interaction, hyperthermophiles, ionic bonds, ionic networks, kinetics, protein folding, protein structure, *Pyrococcus furiosus*, site-directed mutagenesis, *Sulfolobus solfataricus*, temperature optimum, thermal inactivation, thermoactivity, thermodynamics, thermophiles, *Thermoplasma acidophilum*, thermostability.

Contents

Summary

In order to grow optimally at temperatures between 60 ºC and 115 ºC, thermophilic and hyperthermophilic organisms must possess enzymes (*extremozymes*) that are both stable and catalytically active at those temperatures. The comparative molecular enzymology of extremozymes and their mesophilic homologues has generated a wealth of knowledge on the structural basis of enzyme thermostability, and this has been supplemented more recently through laboratory-based evolution experiments. However, thermostability does not guarantee thermoactivity, and the temperature optimum of enzymic activity is often lower than would be expected from the protein's global stability. This paper reviews our current knowledge on thermostability and thermoactivity, and explores the relationship between them.

1. Introduction

Extremophiles are microorganisms that grow optimally in some of Earth's most hostile environments of temperature (–2 ºC to15 ºC and 60 ºC to 115 ºC), salinity (2–5 M NaCl), pH (<4 and >9) and/or pressure. Many extremophiles identified to date are members of the Archaea, although extremophilic Bacteria are known and phylogenetic analyses of environmentally-derived DNA indicate that mesophilic Archaea are also abundant.

To thrive in such environmental extremes, these organisms require cellular components that are naturally resistant to, and functional in, conditions that were once thought incompatible with life. Thus many extremophiles, particularly the thermophiles, are an excellent source of hyperstable macromolecules, and there is no doubt that the discovery of extremophiles, and especially of Archaea, has stimulated a wealth of fundamental and applied research into one macromolecular category in particular, namely extremophilic enzymes. These enzymes, known as *Extremozymes*, may also possess unique catalytic activities and substrate specificities in that the distinct evolutionary lineage of the Archaea, for example, and their adaptation to extreme environments, have lead to their possessing unusual metabolic routes that are catalyzed by enzymes not found in mesophilic organisms. Consequently, it is the novel catalytic ability *and* the remarkable stability of extremozymes that makes them both fascinating and worthy of detailed investigations.

This article will discuss extremozymes from thermophilic (55–80 ºC) and hyperthermophilic (80–115 ºC) organisms, our objectives being to concentrate on two major topics:

- The structural basis of enzyme thermostability
- The nature of thermoactivity and how it relates to thermostability

While such concerns are of fundamental interest to the molecular enzymologist and protein chemist, an understanding of them is also crucial to the enzyme engineer who wishes to tailor their enzymes of choice to meet particular biotechnological needs. As we hope to demonstrate, it is becoming increasingly clear that thermostability may be a prerequisite for thermoactivity, but it does not guarantee it; in fact, although the

connection between the two is in part inseparable, it is not as obvious as once thought. Our approach will be to consider, albeit briefly, theoretical aspects of protein stability and enzyme activity, and then to see how observations on extremozymes, both natural and engineered, support and extend those considerations.

2. Enzyme Stability

2.1. Thermodynamic Stability

If one assumes a two-state model for protein unfolding and folding, where the protein exists only in the folded (F) and unfolded (U) forms, then in a reversible system:

$$F \underset{k_2}{\overset{k_1}{\rightleftharpoons}} U$$

where k_1 and k_2 are the rate constants for the unfolding and folding reactions, respectively. Thus K_{eq} (the equilibrium constant) = [U]/[F] = k_1/k_2, and ΔG_{FU} (the free energy difference between folded and unfolded forms) is given by:

$$\Delta G_{FU} = -R \cdot T \ln K_{eq} \qquad (1)$$

where R = the gas constant and T is the absolute temperature.

ΔG_{FU} is a measure of the *thermodynamic stability* of the protein and is most easily calculated from the determination of K_{eq} values at varying concentrations of a denaturant such as urea or guanidine hydrochloride. At each concentration the degree of unfolding is measured by fluorescence or circular dichroism spectroscopy and then, assuming a linear relationship between ΔG_{FU} and denaturant concentration, the value of ΔG_{FU} in the absence of denaturant can be found by extrapolation. For most proteins, this value of ΔG_{FU} is small (e.g., 20–60 kJ mol^{-1}), equivalent to a few noncovalent interactions, and thus proteins appear to be only marginally stable at their *in vivo* temperatures of operation. However, even though ΔG_{FU} is small in magnitude, it is in fact the combination of two large opposing contributions from the enthalpic (ΔH_{FU}) and entropic (ΔS_{FU}) changes between F and U:

$$\Delta G_{FU} = \Delta H_{FU} - T\Delta S_{FU} \qquad (2)$$

With respect to the *thermo*stability of a protein, it is now necessary to consider how ΔG_{FU} varies with temperature, which in turn depends on the thermal variation of ΔH_{FU} and ΔS_{FU}. In fact, ΔH_{FU} and ΔS_{FU} show relatively large temperature dependencies that are a consequence of the change in heat capacity (ΔC_p) associated with the unfolding of the protein. That is, U has a higher C_p than F, a fact that has been considered to be associated predominantly with the ordering of water molecules when hydrophobic residues are inserted into the solvent, and is thus correlated with changes in solvent accessible surface areas on unfolding of the protein. However, more recently, these ΔC_p effects have been shown to be expected for any system made up of a multiplicity of weak interactions, of which hydrophobic interactions are just a special case. Whatever

the origin of the ΔC_p effects, ΔH_{FU} and ΔS_{FU} change with temperature according to the equations:

$$d\left(\Delta H_{FU}\right)/dT = \Delta C_P \qquad (3)$$

$$d\left(\Delta S_{FU}\right)/dT = T \cdot \Delta C_P \qquad (4)$$

Typical changes in the thermodynamic parameters with temperature for a mesophilic protein are shown schematically in Figure 1, the values being plotted to generate the $\Delta G_{folding}$. Clearly the protein is most stable (the point at which the $\Delta G_{folding}$ is at its most negative value) at ~25 ºC, but it undergoes both thermal and cold denaturation. The temperature at which there are equal quantities of U and F (Keq = 1; $\Delta G_{FU} = 0$) is called the melting temperature (T_m), and this can also be used as a measure of the protein's thermodynamic stability. Figure 1 shows that there are in fact two T_m values, corresponding to cold and thermal denaturation, respectively; however, in the following discussion on protein thermostability, only the higher value will be considered.

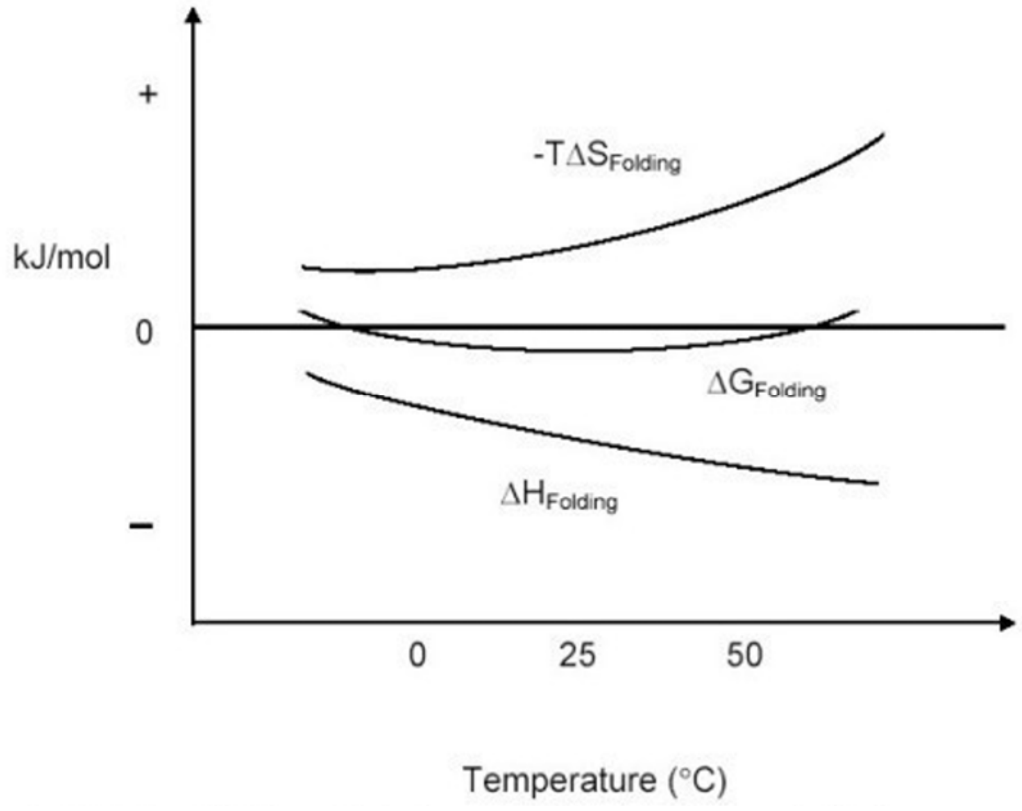

Figure 1. Typical contributions of enthalpy and conformational entropy to the free energy of folding of a protein.
The schematic diagram illustrates how the thermodynamic parameters for protein folding vary with temperature, and shows how the relatively-small value of the free

energy of folding ($\Delta G_{folding}$) of a mesophilic globular protein comprises relatively-large enthalpic ($\Delta H_{folding}$) and entropic ($-T\Delta S_{folding}$) contributions.

As will be evident from the data in Figure 1, increasing the thermostability of a protein must involve a shift in the $\Delta G_{folding}$ versus temperature profile, and there are a number of possible ways in which this can happen: the curve can be shifted to higher temperatures, it can be broadened, or the curve can be shifted to more negative values across the whole temperature range (Figure 2). In all three possible cases, the protein is more thermostable than is the mesophilic protein; that is, at the higher temperatures, the F–U equilibrium position has been shifted towards the folded form (values of ΔG_F are more negative), with a corresponding increase in the value of T_m.

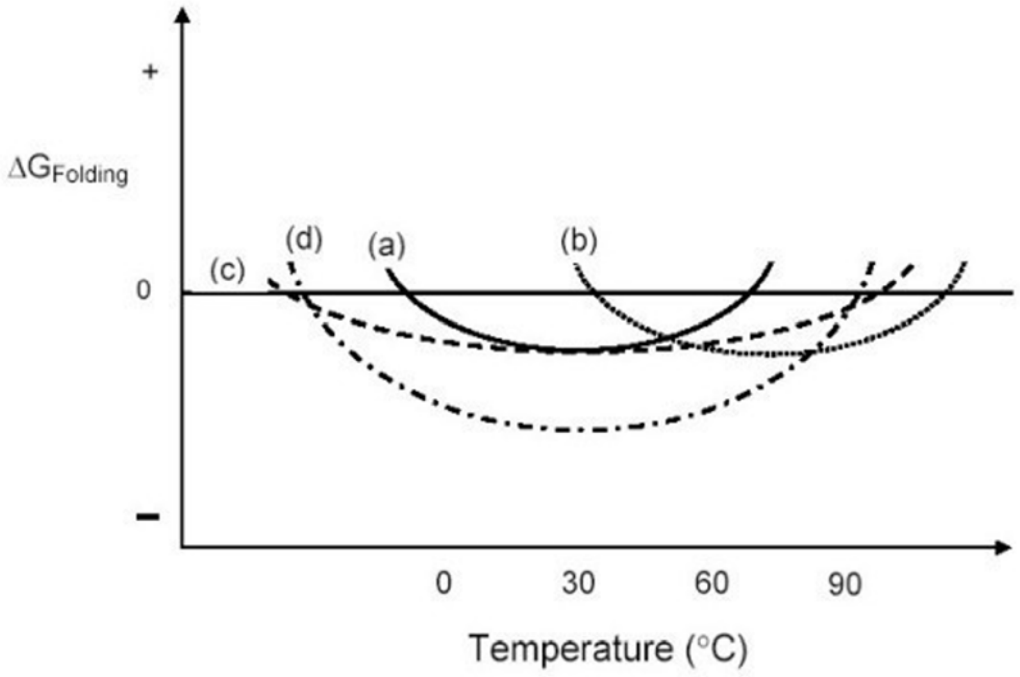

Figure 2. Potential thermodynamic strategies for increasing the thermostability of a protein.

A typical free energy curve ($\Delta G_{folding}$) for a mesophilic enzyme is shown (a), as are also the possible strategies for increasing its thermostability by shifting the free energy curve to higher temperatures (b), broadening the curve (c), or increasing the $\Delta G_{folding}$ at all temperatures (d).

In addition, differential scanning calorimetry (DSC) can be used to determine values of T_m, ΔH_m (ΔH at T_m), and ΔC_p, from which ΔG_{FU} as a function of temperature can then be calculated using the Gibbs-Helmholtz equation:

$$\Delta G_{FU}(T) = \Delta H_m \left[1 - (T/T_m)\right] - \Delta C_P \left[(T_m - T) + T \cdot \ln(T/T_m)\right] \qquad (5)$$

2.2. Kinetic Stability

A practical difficulty with measuring the thermostability of an enzyme is that thermal unfolding is often irreversible, and in such cases the thermodynamic stability cannot be measured as the analysis assumes a reversible process. Therefore, it is common to determine the *rate* of irreversible thermal inactivation of an enzyme as another measure of its thermostability. An extension of the two-state model to take account of the irreversible process might therefore be:

$$F \underset{k_2}{\overset{k_1}{\rightleftarrows}} U \xrightarrow{k_3} X$$

Where X is the irreversibly inactivated state and k_3 is the rate constant describing its formation from U. If the steady-state assumption is made, at least over the period of experimental observation, [U] is constant, and then:

$$k_1[F] = (k_2 + k_3)[U] \tag{6}$$

$$[U] = \frac{k_1[F]}{(k_2 + k_3)} \tag{7}$$

$$\text{Rate of thermoinactivation} = \frac{dX}{dt} = k_3[U] = \frac{k_1 k_3[F]}{(k_2 + k_3)} \tag{8}$$

If $k_3 >> k_2$, as indeed it might be at high temperatures, then:

$$\frac{dX}{dt} = k_1[F] \tag{9}$$

Thus, the observed first-order rate constant for the thermal inactivation of the enzyme (k_{inact}) is equal to k_1, the rate constant for the F to U transition. It should be noted that a similar solution is reached if, instead of assuming the unfolding step to be at equilibrium, the kinetics of the folding and unfolding processes are taken into account, together with the kinetics of the irreversible process.

The value of k_1 is determined by ΔG^* (the activation energy for F to U), in contrast to the thermodynamic stability which is determined by ΔG (Figure 3). It has been proposed that many proteins, and in particular thermophilic proteins, are designed evolutionarily to have significant kinetic stability in addition to their thermodynamic stability, and in the few cases where it has been measured, k_1 is indeed lower than that for mesophilic homologues.

The exact nature of the U to X transition could be an irreversible covalent modification such as hydrolysis or cleavage of the polypeptide chain, deamidation of glutamine and asparagine, or destruction of thermolabile amino acids (e.g., cysteine, methionine,

serine, or threonine). If aggregation of U is the cause of the irreversibility, then the above analysis is more complicated than the simple model described, k_3 no longer being a first-order rate constant.

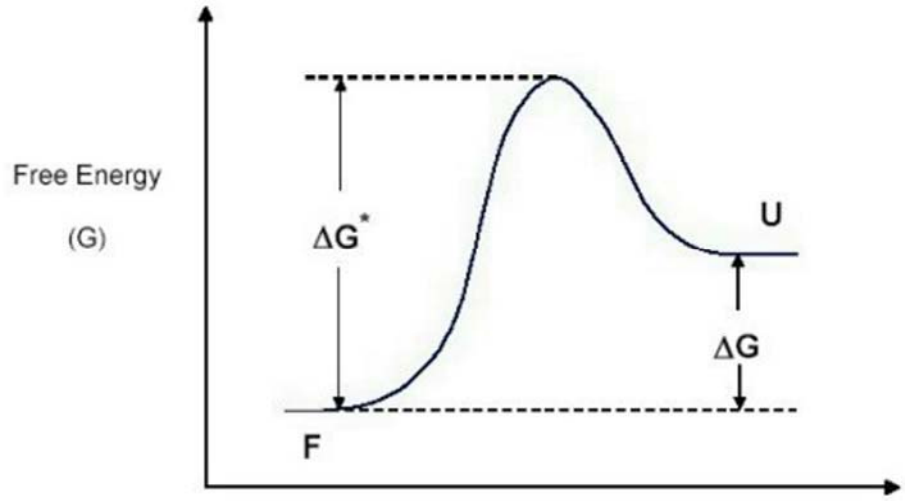

Figure 3. A two-state protein-unfolding coordinate to illustrate thermodynamic and kinetic stability

The schematic diagram illustrates the change in free energy (G^o) along the coordinate describing the unfolding of a folded protein (F) to its unfolded form (U). ΔG is the free energy change between F and U and is a measure of the thermodynamic stability of the protein. ΔG^* is the free energy change between F and the transition state of the process, and may be a measure of the kinetic stability.

2.3. Thermostability

The increased stability of an enzyme compared with its mesophilic homologue can be achieved by a lower rate of unfolding (lower value of k_1 in the above analyses) and/or a higher rate of refolding (increased value of k_2). Both of these effects will serve to decrease ΔG_{FU} (equilibrium is shifted towards F, the folded form of the enzyme), increase T_m, and decrease the rate of thermal inactivation. Thus the enzyme is thermodynamically and kinetically stabilized at high temperatures. Due to the often irreversible nature of thermal inactivation, only a few thermostable proteins have been fully characterized thermodynamically, and these are mostly monomeric, single domain structures that reversibly refold. All three types of shifted $\Delta G_{folding}$ curves (Figure 2) have been observed, and where the individual kinetic constants have been measured, stability appears to be due to decreased rates of unfolding (lowered values of k_1). The question, now, is how is this achieved in structural terms?

3. The Structural Basis of Thermostability

From a consideration of the temperature dependence of the thermodynamic parameters given in Figure 1, thermostabilization of a protein could be achieved in a number of ways. Equally clear, though, is the fact that even hyperthermostable proteins have

relatively small values of ΔG_{FU} at their points of optimal stability, and therefore the structural features contributing to that stability will probably be subtle in nature. Consequently, it is essential to place differences between mesophilic and thermostable proteins in a *structural* context in recognition that thermostability can only be fully understood at the three-dimensional level of protein structure. With the current explosion in genome sequence data, and therefore in the number of protein sequences, it is tempting to make comparisons between mesophilic and thermophilic proteins at the level of amino acid sequence. In our view, the significance of such comparisons is limited, and the value of genome sequencing to our understanding of protein thermostability will only come at the level of *structural genomics*. There is little doubt that within the next few years a massive amount of structural data will soon be available to supplement the considerations outlined in this paper.

In the following discussion, potential stabilizing features are discussed from a theoretical perspective; where appropriate, reference will be made to experimental observations, but only in general terms and where large-scale comparisons have been made. Specific examples will be given in the later sections, when individual thermostable enzymes will be described.

3.1. Conformational Flexibility and Loop Regions

A reduction in the conformational entropy of the unfolded state of a protein will reduce the value of ΔS_{FU} and consequently will promote folding, giving an increase in stability. Consistent with this, comparison of 20 genomic sequences showed that thermophilic proteins are shorter than their mesophilic homologues and, importantly from the point of view of putting this into a structural context, where atomic structures were available the shortening was found to be mainly achieved within the loop regions of the protein or at the N- and C-termini. This is an understandable observation, in that changes to a protein can most easily be accommodated in the loop regions and termini without disturbing the overall fold (and probably the function) of the protein.

Through their inherent flexibility, loops are also considered to be potential initiation points for thermal denaturation, and therefore a reduction in their size (and hence flexibility) may contribute to the protein's kinetic stability. This flexibility may also be reduced by additional stabilizing interactions (enthalpic contributions to ΔG_{FU}) and, as discussed later, ionic bonds may have a particular role to play in this context.

-
-
-

Bibliography

Adams M.W.W. and Kelly R.M. (Eds.) (2001). *Methods in Enzymology.* Volumes 330, 331, and 334. Hyperthermophilic Enzymes, Parts A, B, and C. [A comprehensive three-volume series on thermophiles, hyperthermophiles, and their enzymes. The papers cover both fundamental and applied enzymology.]

Cooper A. (1999). Thermodynamics of protein folding and stability. *Protein: a Comprehensive Treatise* **2**, 217–270. [An excellent account of basic thermodynamic principles relevant to protein folding and their experimental analysis.]

Daniel R.M., Danson M.J., and Eisenthal R. (2001). The temperature optima of enzymes: new perspectives on an old phenomenon. *Trends in Biochemical Sciences* **26**, 223–225. [This paper reassesses the concept of true temperature optima for enzymes and relates closely to the relationship between enzyme structure, stability and catalytic activity]

Bell G.S., Russell R.J.M., Connaris H., Hough D.W., Danson M.J.and Taylor G.L. (2002). Stepwise adaptations of citrate synthase to survival at life's extremes. *Eur. J. Biochem.* **269**, 6250-6260. [This paper summarizes the work on the structure-stability analyses of citrate synthases from thermophilic and psychrophilic organisms.]

Dill K.A. (1990). Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155. [This excellent paper is a comprehensive review on the multiplicity of forces involved in proteins attaining their three-dimensional structures.]

Hough D.W. and Danson M.J. (1999). Extremozymes. *Current Opinion in Chemical Biology* **3**, 39–46. [This review focuses on the new opportunities that extremozymes offer for biocatalysis and biotransformations]

Jaenicke R. and Bohme G. (1999). The stability of proteins in extreme environments. *Curent Opinion in Strucural Biology* **8**, 738–748. [This is an excellent review from one of the major contributing research groups in extremozyme stability.]

Karshikoff A. and Ladenstein R. (2001). Ion pairs and the thermotolerance of proteins from hyperthermophiles: a traffic rule for hot roads. *Trends in Biochemical Sciences* **26**, 550–556.[This paper gives a precise and readable account of ion pairs and thermostability.]

Kumar S., Tsai C-J., and Nussinov R. (2001). Thermodynamic differences among homologous thermophilic and mesophilic proteins. *Biochemistry USA* **40**, 14152–14165. [An analysis of experimental protein stability curves, with a discussion on the correlation of thermodynamic parameters with protein and water structure.]

Makhatadze G.I. and Privalov P.L. (1995). Energetics of protein structure. *Advances in Protein Chemistry* **47**, 307–425. [A comprehensive treatise on experimental and theoretical aspects of protein folding.]

Sterner R. and Liebl W. (2001). Thermophilic adaptation of proteins. *Critical Reviews in Biochemistry and Molecular Biology* **36**, 39–106. [This paper is a comprehensive review of thermostable proteins. It is very well referenced.]

Thompson M.J. and Eisenberg D. (1999). Transproteomic evidence of a loop-deletion mechanism for enhancing protein stability. *Journal of Molecular Biology* **290**, 595–604. [The analysis of genome sequence data to correlate loop size with protein thermostability.]

Vielle C. and Zeikus G.J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms of thermostability. *Microbiology and Molecular Biology Reviews* **65**, 1–43. [A comprehensive review of thermostable enzymes. An excellent source of references to the relevant original literature.]

Wintrode P.L. and Arnold F.H. (2001). Temperature adaptation of enzymes: lessons from laboratory evolution. *Advances in Protein Chemistry* **55**, 161–225. [A lucid and innovative account of the

contribution that laboratory evolution has made to the understanding of the structure, function and stability of thermophilic enzymes.]

Biographical Sketches

Michael Danson is currently a Professor of Biochemistry and Director of the Centre for Extremophile Research in the Department of Biology and Biochemistry at the University of Bath, UK. He received his BSc Honors Degree in Biological Sciences in 1970 from the University of Leicester, UK, where he then read for his PhD in the area of molecular enzymology under the supervision of Professor David Weitzman in the Department of Biochemistry. In 1974, he joined the group of Professor Richard Perham in Biochemistry at the University of Cambridge, where he worked for five years on the enzymology and protein chemistry of the pyruvate dehydrogenase multienzyme complex. In 1979, he was appointed to the position of lecturer in Biochemistry at the University of Bath where in 1981 he began his work on the enzymology of extremophiles. His research is centered on the structural basis of enzyme stability and function to extremes of temperature and salinity, and in the last few years has also included the biotechnological applications of extremozymes, and the isolation and cultivation of novel extremophiles. In 1996, he and colleagues launched the Centre for Extremophile Research, and he was promoted to a Chair in Biochemistry in 1997. He has spent two Sabbatical leaves, firstly in the University of Calgary, Canada, and then in the University of Waikato, New Zealand, and is the author of 120 original publications.

David Hough is a Reader in Biochemistry in the Department of Biology and Biochemistry at the University of Bath, UK, and a founder member of the Centre for Extremophile Research. He obtained a BSc in Biochemistry from the University of Manchester Institute of Science and Technology in 1966 and his DPhil was obtained in 1970 from the School of Biological Sciences, University of Sussex, UK. After spending a year as a research fellow at ENSC Montpellier, France, David joined Prof. George Stevenson's group at the Tenovus Research Laboratory, University of Southampton, UK, where he spent six years. In 1976, he was appointed as Lecturer in Biochemistry at the University of Bath, UK, where he has worked on the structure, stability and activity of extremozymes and, more recently, has focussed on their biotechnological applications. He has collaborated with a number of European research groups working on extremophiles and is the author of 120 original publications.