Biophysical Chemistry 29 (1988) 181–193
Elsevier

BPC 01227

Structure and stability of thermophilic enzymes

Studies on thermolysin *

Angelo Fontana

Department of Organic Chemistry, Biopolymer Research Centre of CNR, University of Padua, Padua, Italy

Accepted 15 October 1987

Thermophilic enzyme; Thermolysin; Limited proteolysis; Protein domain; Protein dynamics

The molecular mechanisms responsible for the unusual stability of enzymes isolated from thermophilic microorganisms are much more complex and subtle than was originally thought. In particular, a general mechanism cannot be proposed, since individual enzymes can be stabilized by specific molecular interactions and forces. The results of studies on thermophilic enzymes obtained in recent years in our laboratory will be summarized, with particular emphasis being placed on those obtained with thermolysin, a stable metalloendopeptidase isolated from *Bacillus thermoproteolyticus*. Fragmentation of thermolysin by limited proteolysis by added protease (subtilisin) or autolysis mediated by heat or the ion-chelating agent EDTA leads to quite selective peptide bond fissions, allowing isolation of 'nicked' thermolysin species. Correlation of the sites of proteolytic cleavage with the known three-dimensional structure of thermolysin allowed us to infer some of the key characteristics of the structure, folding, dynamics and stability of the thermolysin molecule. The potential utility of these and other studies on thermophilic enzymes in devising strategies for enhancing the stability of mesophilic enzymes using genetic engineering techniques is discussed.

1. Introduction

Much attention in research has been devoted in recent years to enzymes and proteins isolated from thermophilic microorganisms growing optimally at 60–90 °C [1–7]. These numerous studies on the functional and molecular properties of thermophilic enzymes have clearly established that these enzymes are generally much more resistant to heat and most common protein denaturants than their counterparts from mesophilic sources [8]. The interest of biochemists has so far been focused primarily on the structure-function-stability properties of thermophilic enzymes, with the general aim of determining the molecular mechanisms responsible for their unusual stability. More recently, since enzyme stability is one important parameter in the field of enzyme technology [9], potentialities and advantages of thermophilic enzymes for practical applications are becoming more clearly recognized [10–15].

In this article, some general properties and characteristics of thermophilic enzymes are discussed, with special emphasis on the structure-stability-folding-dynamic properties of thermolysin, a stable metalloendopeptidase isolated from *Bacillus thermoproteolyticus* [16,17]. There is a vast amount of literature available on the isolation and characterization of thermophilic enzymes, with articles covering a wide range of aspects and inter-

* Presented at the 9th International Biophysics Congress, Jerusalem, August 23–28, 1987.

Correspondence address: A. Fontana, Department of Organic Chemistry, Biopolymer Research Centre of CNR, Via Marzolo 1, 35131 Padua, Italy.

Abbreviations: nicked thermolysin, the autolysis product of thermolysin composed of fragments associated in a stable complex; FAGLA, N_{α}-furylacryloylglycyl-L-leucine amide; HPLC, high-performance liquid chromatography.

ests, and therefore the reader may find herewith a somewhat personal selection of issues covered and literature references. More detailed accounts on the various functional and molecular aspects of thermophilic enzymes (and microorganisms) are available [1–7,10–15].

2. Molecular mechanisms of thermostability

Thermophilic enzymes are not at all unique with respect to their structural properties, but are 'typical' proteins characterized by an intrinsic stability, even if in some cases, e.g., α-amylase [18] and thermolysin [19–21], protein thermal stability can be achieved extrinsically by addition of suitable effectors (metal ions, coenzymes, etc.). All current evidence indicates that the enhanced stability of thermophilic enzymes cannot be attributed to a common determinant, but is the result of a variety of stabilizing effects including hydrophobic interactions, ionic and hydrogen bonding, disulfide bonds, metal binding, etc., much the same as has already been observed with mesophilic proteins. A quantitative analysis of all these factors involved in protein stability appears to be very complicated, since a large number of stabilizing effects must be considered. Considering that the net free energy of stabilization of a protein (mesophilic and thermophilic) is numerically quite small, usually of the order of 5–15 kcal/mol and that the stabilization provided by a single hydrogen bond or salt bridge is of the order of 1–3 kcal/mol [22–25], it is clear that interaction energies of this magnitude can arise in different ways through appropriate combinations of the weak bonds commonly involved in stabilizing proteins. On this basis, there is little hope that a universal molecular mechanism of thermostability can be found, since different proteins may be stabilized in different ways.

Initial efforts of explaining thermostability were carried out by comparing the amino acid composition of two groups of proteins, one from thermophilic and the other from mesophilic organisms. Hydrophobicity has been believed to play a role in thermostabilization, considering that the strength of hydrophobic interactions increases as a function of temperature, at least up to 60–85°C [26]. Thus, it was anticipated that thermophilic proteins were, in general, more hydrophobic than corresponding mesophilic proteins. A good correlation has been found between certain macroscopic parameters calculated from the amino acid composition of approx. 20 mesophilic vs. thermophilic proteins from closely related species [27]. These were the hydrophobic index, the ratio of polar to nonpolar residues, the Arg/(Arg + Lys) ratio, plus the % potential α-helix or β-sheet (α-index or β-index). The aliphatic index (the relative volume of a protein occupied by residues with aliphatic side chains: Ala, Val, Ile and Leu) also appears to be related to thermal stability [28]. A parameter related to the average residue volume and hydrophobicity also correlates with the transition temperature of 14 nonthermophilic proteins [29]. However, there efforts to relate intrinsic thermostability to hydrophobicity do not seem to be convincing [30].

A much more fruitful approach has been a detailed examination of sets of homologous proteins from thermophilic and mesophilic sources in terms of amino acid sequences and three-dimensional structures. Perutz and Raidt [31] compared the structures of various clostridial ferredoxins from thermophilic and mesophilic sources and were able to attribute thermostability mainly to the formation of a few additional salt bridges in the thermophilic ferredoxin. Walker et al. [32–34] observed that the *B. stearothermophilus* glyceraldehyde-3-phosphate dehydrogenase contains more arginines than the lobster enzyme and that two of these extra arginines form salt bridges across the subunit interface stabilizing the tetrameric structure of the thermophilic enzyme. The dehydrogenase from the extreme thermophile *Thermus aquaticus*, which is even more thermostable, shows a greater number of intersubunit and surface salt bridges; however, with this enzyme it was recognized that thermal stabilization is achieved through both ionic and hydrophobic interactions [34]. Argos et al. [35], on the basis of a comparative analysis of sequences and structures of the thermophilic and mesophilic molecules of ferredoxin, glyceraldehyde-3-phosphate dehydrogenase and lactate dehydrogenase, proposed that

thermal stability can be achieved by the addition of many small changes throughout the molecule without significant alterations in the backbone conformation. These changes serve to increase the external hydrophobicity, stabilize helices and increase internal residue packing arrangements. It was observed that thermophilic enzymes are characterized by a significant preference for Ala, Thr and Arg.

The most clearcut demonstration that *subtle* structural differences between a thermophilic and a mesophilic protein are sufficient to account for their enhanced stability came from the analysis of a series of temperature-sensitive mutants of phage lysozyme [36–40], β-galactosidase from *E. coli* [41] and the α-subunit of tryptophan synthetase [42]. These studies showed that a substantial effect in thermal stability of a protein can be achieved simply by a *single* amino acid residue substitution in the protein polypeptide chain. Matthews and co-workers [36] showed that replacement of an arginine by a histidine in lysozyme from the bacteriophage T4 does not affect the three-dimensional structure of the protein but lowers its melting temperature by about 14°C. More recently, many other mutants of T4 phage lysozyme that differ from the wild type in stability have been obtained and the detailed structural and physicochemical analysis of these mutants is in progress [37–40]. Several other protein systems differing in a single amino acid residue and stability have been described [43–46].

3. Rigidity and catalytic efficiency of thermophilic enzymes

It is known that the static picture of globular proteins derived from crystallographic analyses represents an average structure, since proteins are dynamic molecules characterized by fluctuations of specific loops of the polypeptide chain, as clearly emphasized in recent years by numerous studies employing a number of physicochemical techniques and theoretical calculations [47–50]. It has been found that thermostable proteins from thermophiles are quite rigid molecules at room temperature with respect to their mesophilic counterparts, as a direct consequence of the 'clamping' effect of protein structure brought about by the interactions and forces stabilizing thermophilic proteins. The rigidity of these proteins was clearly established using proteolytic enzymes as probes of structure. In fact, it is a general notion that proteases attack native globular proteins only if specific chain loops are characterized by availability at the protein surface and at the same time if they show the proper degree of adaptability at the active site of the protease (see also below). Several thermophilic enzymes, including 6-phosphogluconate dehydrogenase from *B. stearothermophilus* [8], asparaginase from a *Thermus* species [51] and a glucosidase from *B. thermoglucosidicus* [52], were found much more resistant to the action of several proteolytic enzymes with respect to mesophilic counterparts. The results of proteolysis experiments, together with those obtained using spectroscopic techniques [53–55], indicate that there is an inverse correlation between flexibility (motility) and thermal stability of protein molecules.

The rigidity of thermophilic enzymes has an adverse effect on their catalytic efficiency, since the appropriate degree of flexibility is required for enzyme catalysis [47–50,56,57]. In fact, considering that the rates of enzyme reactions normally double on being performed at a temperature 10°C higher, one would expect that thermophilic enzymes should have an extremely high activity at thermophilic temperatures. In contrast, it has been observed that the specific activity of thermophilic enzymes is considerably less than would be predicted from the specific activity of their mesophilic counterparts [3,58]. Consequently, since a thermophilic enzyme is less effective at lower temperatures, it seems that an increase in thermostability brings about a reduction in catalytic efficiency, and vice versa. The reduced efficiency of thermophilic enzymes is compensated by the higher temperature of their habitat, so that similar specific activities are observed for both thermophilic and mesophilic enzymes at their respective temperature optima. Only at the normal temperatures of the organism from which they are derived are thermophilic enzymes sufficiently flexible to be fully active and yet rigid enough not to be denatured [30]. This low catalytic efficiency of thermo-

philic enzymes at room temperature has often been observed [3,30,58]. A strict inverse correlation between specific activity and optimum of activity (i.e., thermostability) has been observed in the case of enolase from mesophilic (rabbit and yeast) and thermophilic (*T. aquaticus* and *Thermus X-1*) sources [59] and, more recently, with four mutants of kanamycin nucleotidyltransferase [44].

4. Thermolysin

Thermolysin is an exceptionally thermostable neutral metalloendopeptidase of molecular mass 34.6 kDa isolated from *B. thermoproteolyticus* [16,17]. This enzyme is constituted by a single polypeptide chain of 316 amino acid residues, lacking thiol or disulfide bonds, containing a catalytically essential Zn^{2+} and four calcium ions [60–64]. Kinetic studies have shown that thermolysin possesses quite broad substrate specificity, hydrolyzing preferentially peptide bonds on the amino side of hydrophobic amino acid residues [64]. Because of its unusual stability to heat, thermolysin was a suitable molecule for carrying out structural studies to elucidate the molecular basis of thermostability of thermophilic enzymes. To this end, the three-dimensional structure of thermolysin was determined by X-ray methods [65–67], this being the first complete structure determination of a protein from a thermophilic source. This structural analysis allowed Matthews and co-workers [65–68] to conclude, for the first time, that thermophilic enzymes do not possess special structural features compared to their mesophilic counterparts and that, in the case of thermolysin, ionic interactions and calcium binding appear to play a major role in determining protein stability. Recently, the structure of thermolysin has been refined to a nominal resolution of 1.6 Å [69] and the mode of binding of a number of inhibitors has been determined by X-ray crystallography [70–73]. Thus, thermolysin can be listed among the most well characterized enzyme molecules, making this protein an interesting model for studying structure-function relationships as well as folding and stability properties of proteins.

5. Domains and subdomains in thermolysin

The wealth of structural data of proteins currently available allows a description of the structure of a globular protein in a hierarchic fashion, ranging from elements of secondary structure (helices, sheets) to subdomains, domains and whole protein [74–78]. The word domain of a protein is currently used to indicate large subassemblies of secondary structure elements which appear tightly packed in the crystal structure of globular proteins. Protein domains are almost invariably seen in globular proteins with more than about 100 amino acids and such domains often appear rather well separated form each other, so that protein molecules appear to be constituted of lobes. Thermolysin shows a quite peculiar bilobal morphology, with two distinct domains of equal size (residues 1–157 and 158–316) and the active site located at the interface between them [65,74]. The NH_2-terminal domain is mainly β-pleated sheet

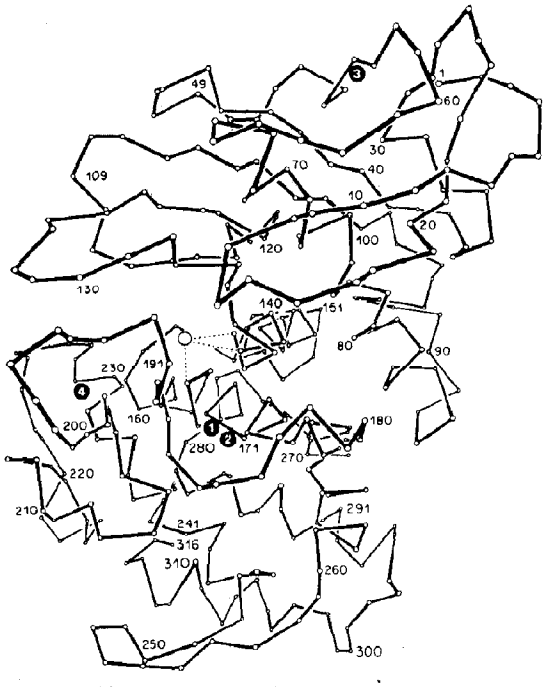

Fig. 1. Backbone diagram of the thermolysin crystallographic structure. The zinc (○) and calcium ions (●) are also indicated (taken from ref. 66).

and the COOH-terminal one mainly α-helical (see fig. 1).

Wetlaufer [74] was the first to emphasize the structural role of domains in globular proteins and to propose that domains could represent intermediates in the folding process of protein molecules. It is conceivable to suggest that specific segments of an unfolded polypeptide chain first refold to individual domains, which then associate and interact to give the final tertiary structure, much the same as do subunits in oligomeric proteins. The major implication of this model of protein folding by a mechanism of modular assembly is that isolated protein fragments corresponding to domains in the intact protein are expected to be able to fold into a native-like structure independently from the rest of the polypeptide chain, resembling in their properties a low-molecular-mass globular protein [79]. Indeed, this possibility was tested experimentally in our laboratory using fragments of thermolysin obtained by cyanogen bromide cleavage of the protein at the level of the two methionine residues in positions 120 and 205 of the polypeptide chain of 316 amino acid residues [80–83]. Fragments 1–120, 121–205 and 206–316 were thus prepared [80]. By using a small amount of reagent and shorter reaction times it was also possible to prepare the 'overlapping' fragments 1–205 and 121–316 [84,85]. All these fragments have been isolated to homogeneity and their conformational properties investigated [80,83]. Overall, the results obtained established that protein fragments corresponding to domains, as well as subdomains (folding units or supersecondary structures), can refold into a native-like and stable structure.

The most detailed studies were carried out on fragment 121–316 [84], comprising entirely the 'all-α' COOH-terminal domain 158–316, as well as fragment 206–316 [82,83]. Both fragments are able to refold in aqueous solution at neutral pH into a structure with native-like characteristics, as judged from circular dichroism (CD) measurements and immunochemical properties using rabbit anti-thermolysin antibodies [83,84]. The values for the α-helix content of fragments 121–316 and 206–316, calculated from far-ultraviolet CD spectra [86,87], were in essential agreement with those expected for a native-like structure of the fragments and calculated from the known three-dimensional structure of thermolysin. Furthermore, the recognition and precipitation of anti-thermolysin antibodies observed with both fragments are indicative of a close structural relationship between fragments and intact native thermolysin [80,81,88]. In fact, antibodies elicited towards a globular protein are specific for antigenic determinants which are located in the more exposed regions of the protein molecule and thus able to probe similarities of important details (loops, corners) of the three-dimensional structure in the isolated fragments and the native parent protein [89,90].

In order to establish the minimum size for a COOH-terminal fragment of thermolysin capable of independent folding, fragment 206–316 was subjected to limited proteolysis by subtilisin [91,92]. The main component of the proteolytic mixture was a fragment of approx. 6 kDa which was purified by gel filtration and shown to correspond to sequence 255–316. Also fragment 255–316 was shown to be capable of independent, native-like folding in aqueous solution at neutral pH, signifying that it is possible to isolate stable supersecondary structures (subdomains or folding units) from globular proteins.

Protein fragments possessing domain characteristics are expected to behave like low-molecular-mass globular proteins, i.e., to show cooperativity of unfolding transitions. Indeed, even the smaller folded fragment 255–316 [91,92] shows cooperative and reversible thermal unfolding. The thermal stability of this fragment (T_m 65 °C) appears quite remarkable when one considers its small size and the fact that it is a simple polypeptide chain of 62 amino acid residues, lacking disulfide bridges, cofactors and strongly bound ions, all characteristics well known to contribute significantly to the folding and stability of protein structures.

It is of interest to relate the folding properties of COOH-terminal fragments of thermolysin to the predicted locations of subdomains in the COOH-terminal structural domain (residues 158–316) of thermolysin using computer algorithms [78,93,94]. These predictions, based on the X-ray structure of thermolysin, all identified a subdo-

main comprising roughly residues 240–316 [78,93], constituted by the assembly of the three COOH-terminal helices (see fig. 1). Rashin [94] computed the location of *stable* subdomains with native-like conformation in the COOH-terminal portion of the thermolysin molecule on the basis of surface area measurements and predicted that subfragments 214–316, 235–316, 255–316, and perhaps also 285–316, should have a good chance of showing independent folding. Overall, the folding properties of the COOH-terminal fragments of thermolysin are in essential agreement with these predictions.

6. Limited proteolysis and autolysis of thermolysin

In analogy to all enzymatic reactions, the proteolytic cleavage of a polypeptide chain occurs only if the site of cleavage can bind and adapt itself in a specific way to the stereochemistry of the active site of the protease. This is difficult to achieve with native globular proteins, whereas denatured proteins are much more susceptible to proteolysis [95]. In a number of cases, an extraordinary lability to enzymatic hydrolysis of a very small number of specific bonds in a native globular protein has been observed and this selective peptide bond fission has been termed 'limited proteolysis' [96]. It is conceivable to suggest that the sites of limited proteolysis in a native globular protein are dictated solely by the stereochemistry of the protein substrate, if a protease of low specificity is employed. In addition, some motility of the substrate protein at the site of cleavage would be required for a proper adaptation to the active site of the protease [97,98].

With the aim of probing the structure and dynamics of thermolysin using the limited proteolysis approach and, hopefully, to prepare nicked thermolysin species of structural and functional interest, we have carried out a number of experiments of limited proteolysis of thermolysin under different experimental conditions and by addition of an external protease (subtilisin) [99] or by autolysis mediated by heat [100] or EDTA [101]. Indeed, the pattern of protein fragmentation thus obtained allowed one to infer some molecular aspects of the protein-protein recognition processes underlying the proteolytic event [102,103].

6.1. Thermolysin S

Incubation of thermolysin at pH 9–10 in the presence of 10 mM $CaCl_2$ for 2 days at room temperature with subtilisin at a molar ratio of 50 : 1 leads to a derivative possessing lower (~ 3%) but intrinsic catalytic activity [99]. This derivative, called thermolysin S, was isolated by an affinity chromatographic step on Sepharose-Gly-D-Phe [104]. It was found that thermolysin S results from a tight association of two polypeptide fragments of apparent molecular masses 24 and 10 kDa. Dissociation of the complex was achieved under strong denaturing conditions, such as gel filtration on a column equilibrated and eluted with 5 M guanidine hydrochloride. The positions of the clip sites were defined by amino acid analysis, end-group determination and amino acid sequencing of the isolated fragments and shown to lie between Thr-4 and Ser-5, between Thr-224 and Gln-225, and also between Gln-225 and Asp-226. Thermolysin S, which is therefore a stable complex of fragments 5–224(225) and 225(226)–316, shows a shift in optimum pH of about 1 unit towards the acid range with respect to intact thermolysin and a K_m that remains essentially unchanged with FAGLA as substrate. Inhibitors of thermolysin such as ethoxyformic anhydride and Zn^{2+} also inactivate the nicked enzyme. The overall conformational properties of thermolysin S appear very similar to those of the parent native enzyme, as judged by CD measurements and by the fact that rabbit antiserum against native thermolysin recognizes and precipitates thermolysin S, as detected by immunodiffusion. The lower activity of thermolysin S with respect to thermolysin can be related to the fact that cleavage occurs near the Asp^{226} residue interacting with His^{231}, involved in the catalytic function of the enzyme [105].

6.2. EDTA-mediated autolysis

X-ray crystallographic studies [65–67] have demonstrated the presence of the functional Zn^{2+} and four calcium-binding sites in the thermolysin

molecule, in agreement with solution studies (ref. 20 and references cited therein). Two of the four calcium ions, Ca(1) and Ca(2), form a double site in which the interatomic distance between them is only 3.8 Å; the two single sites, Ca(3) and Ca(4), involve binding at Asp^{57} and Asp^{200}, respectively [20,65–67] (see fig. 1). The role of calcium in both the thermal denaturation [62] and autolysis [19] of thermolysin has been the subject of intensive investigation [20]. However, the detailed molecular understanding of this role is complicated by the presence of four binding sites and by the fact that calcium ions can stabilize the overall three-dimensional structure of the thermolysin molecule and/or prevent autolytic degradation of the enzyme. Characterization of the pattern of fragmentation (autolysis) of thermolysin in the presence of the metal-chelating agent EDTA allowed some key indications to be obtained about the molecular mechanism of stabilization of thermolysin by calcium ions [101].

Incubation of thermolysin at pH 7.2 in the presence of 10 mM EDTA and 1.5 mM $CaCl_2$ produces fast enzyme inactivation, as a result of almost quantitative autolysis. The nicked protein is a folded species, composed of three tightly associated protein fragments and which can be isolated by gel-filtration chromatography [101]. Dissociation of this complex and isolation of the individual fragments can be achieved under denaturing conditions, such as gel filtration on a column equilibrated with 5 M guanidine hydrochloride or reverse-phase HPLC under acidic conditions. Amino acid composition and sequence analyses of these fragments established that the nicked protein is composed of fragments 1–196, 197–204 and 205–316. In the presence of a lower (1 mM instead of 10 mM) EDTA concentration, it was observed that thermolysin is additionally cleaved at peptide bonds 129–130 and 187–188 [102].

The rapid autolysis of thermolysin in the presence of EDTA at neutral pH can be viewed in terms of removal of the calcium ion(s) from the metalloenzyme leading to some changes in protein structure and of removal of the functional Zn^{2+} leading to enzyme inactivation. At the initial stages of the chelation reaction, the remaining active enzyme could attack a calcium-depleted, folded enzyme species, but highly susceptible to proteolytic attack. The selective cleavage of the peptide bonds Gly^{196}–Ile^{197} and Ser^{204}–Met^{205} by autolysis of thermolysin when dissolved in the presence of 10 mM EDTA implies that the corresponding chain loop 195–205 becomes exposed and flexible in the presence of the metal chelating agent. This proposal appears to be rather well founded, considering previous data on the binding affinity for calcium ions of thermolysin. Matthews and coworkers [68] determined the relative binding affinities of the four calcium ions in thermolysin by soaking protein crystals in the presence of different concentrations of EDTA and subsequently identifying the calcium ions being removed from the protein by this procedure using difference Fourier analysis. It was found, for example, that thermolysin crystals lose two calcium ions in the presence of 5 mM $CaCl_2$ and 5 mM EDTA. Interestingly, complete removal of all four calcium ions could not be achieved even in the presence of an excess of chelating agent. The results of these experiments allowed Matthews et al. [68] to propose the following ranking of binding affinities for the four calcium ions: $Ca(1) \gg Ca(3) > Ca(4) \geq Ca(2)$. The results of the present study are in line with these previous findings, since they demonstrate that fast and selective autolysis occurs at the chain loop 190–205 involved in binding Ca(4) in native thermolysin; this ion is bound to the backbone carbonyl of oxygen of Tyr^{193}, Thr^{194} and Ile^{197}, the hydroxyl group of Thr^{194}, the carboxyl group of Asp^{200} and two water molecules. Actually, Voordouw and Roche [20,106] studied calcium binding of thermolysin in solution and found that at free calcium concentrations of about 5×10^{-5} M two calcium ions dissociated simultaneously from the enzyme and argued that they were the Ca(1) and Ca(2) ions of the double site. On the other hand, current data identify Ca(4) in the earlier events of ion dissociation from the protein and substantiate the proposal advanced by Matthews et al. [67,68] that the function of Ca(4) could be to protect the corresponding binding loop against autolysis. The binding site of Ca(4) is within the most extended region of irregular conformation in the thermolysin molecule, so that, in

the absence of calcium, this portion of the molecule becomes more flexible and a more suitable site of autolytic attack.

The EDTA-mediated fragmentation of thermolysin identifies, among the four protein-bound calcium ions, Ca(4) as a major stabilizer of the thermolysin molecule against autolysis. In this context, it is of interest to recall that neutral protease from *B. subtilis* was found by amino acid sequence analysis to be homologous to thermolysin and to share many properties in common, including ion binding, similar specificity, molecular mass, and even mechanism of action [105,107,108]. Many of the ligands associated with zinc and calcium binding are present in both proteins at identical loci of the polypeptide chain. However, it has been found that the *B. subtilis* protease binds fewer calcium ions than thermolysin, in agreement with the fact that the binding sites for Ca(4) cannot exist in neutral protease due to the exchange of Asp^{200} in thermolysin with Pro^{200} in neutral protease and the deletion of the other three residues of the chelating loop [107]. The major difference between the two proteases lies in their thermal stability, since neutral protease is inactivated at a temperature of about 25 °C below that of thermolysin [108]. Thus, the results of sequence analysis and those of EDTA-mediated autolysis seem to indicate that Ca(4) of thermolysin plays a major role in stabilizing thermolysin with respect to *B. subtilis* neutral protease.

6.3. Thermal autolysis

Thermolysin can be induced to autolyze almost quantitatively when dissolved at a relatively high protein concentration (1 mg/ml) in 50 mM Tris-HCl buffer, containing 10 mM $CaCl_2$, pH 9.0, and heated at 55 °C for 48 h [100,102]. Analysis of the reaction mixture by SDS-polyacrylamide gel electrophoresis showed that the fragmentation pattern is quite specific, since only a few protein fragments of 24, 17, 10 and 7 kDa are visible in the gel. When an aliquot of the proteolytic mixture was applied to a Sephadex G-75 column, eluted with 10 mM Tris-HCl buffer, pH 9.0, containing 10 mM $CaCl_2$, two peaks of protein material were eluted from the column. The first peak contained an aggregated (most likely dimeric), three-fragment complex constituted by fragments of 17, 10 and 7 kDa, while the second contained, besides these fragments, an additional fragment of 24 kDa and also some intact thermolysin. The individual fragments produced by thermal autolysis of thermolysin were isolated to homogeneity by HPLC and their identity established by amino acid analysis after acid hydrolysis and sequencing. Comparison of all these data with the known amino acid sequence of thermolysin [60] allowed unambiguous identification of fragments 1–221, 1–154(155), 155(156)–221 and 224–316 as those produced (and isolated) by thermal autolysis of thermolysin. Interestingly, reverse-phase HPLC allowed separation of fragment 1–154 from fragment 1–155, i.e., this technique was efficient in separating peptide components differing in size by a single amino acid residue, Leu^{155} (unpublished results).

6.4. Correlation of sites of proteolysis with the three-dimensional structure of thermolysin

Fig. 2 shows schematically the sites of cleavage in the polypeptide chain of thermolysin observed by proteolysis with subtilisin or by autolysis mediated by heat or EDTA. The selectivity with which limited proteolysis occurs in the thermolysin molecule, as indicated by the small number of fragments generated, appears to be quite striking. Considering the broad specificity of both subtilisin [109] and thermolysin [64], conformational features of the globular protein substrate clearly must dictate the sites of attack on the protein surface. A careful inspection of the three-dimensional structure of intact thermolysin (fig. 1) provides evidence that the observed sites of cleavage of the thermolysin structure are located at exposed loops (turns or bends) of the thermolysin polypeptide chain, never within segments of ordered secondary structure. For example, thermal autolysis leads to fission at peptide bonds 154–155 and 155–156, located on a loop connecting two helical segments, and also at loop 220–226. Subtilisin also cleaves at this loop, as well as at the amino-terminus. In the presence of EDTA, which removes looser calcium ion(s) from the metalloprotein [19,68], fission oc-

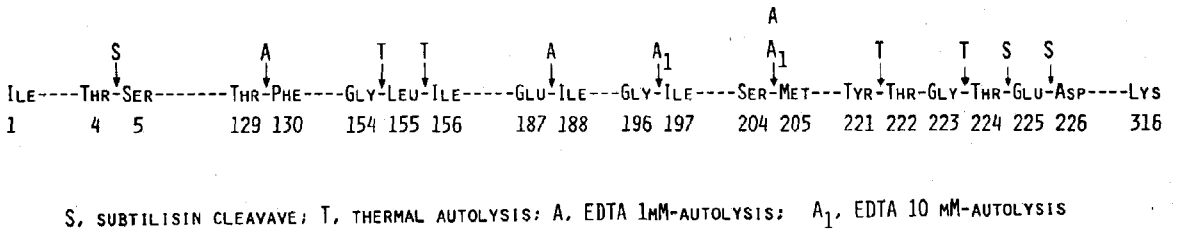

Fig. 2. Schematic representation of the amino acid sequences at the sites of limited proteolysis by subtilisin (S) and autolysis by heat (T) or in the presence of EDTA (A and A_1). Details of the experimental conditions employed to cleave thermolysin by subtilisin and to isolate the nicked species thermolysin S have been reported [99]. Autolysis of thermolysin upon heating [100] or in the presence of 10 mM EDTA [101] or 1 mM EDTA [102] leads to quite selective peptide bond fissions and formation of thermolysin nicked species constituted by two as well as three fragments associated in stable complexes. The sites of cleavage have been determined by isolation and characterization in terms of amino acid composition and sequences of all the thermolysin fragments produced by proteolysis and comparing these data with the known amino acid sequence of thermolysin [60]. S, site of cleavage of the thermolysin polypeptide chain by subtilisin; A, cleavage by autolysis in the presence of 1.5 mM $CaCl_2$ and 1 mM EDTA; A_1, cleavage by autolysis in the presence of 1.5 mM $CaCl_2$ and 10 mM EDTA; T, cleavage by thermal autolysis.

curs at peptide bond 129–130, located on a surface turn, and at a long surface loop 180–210, held in place by calcium ions in the native molecule [20]. Thus, exposed loops of the thermolysin molecule seem to be the loci on the protein structure more easily recognised by the protease active site.

The loops which are easily attacked and cleaved by the protease are characterized by high segmen-

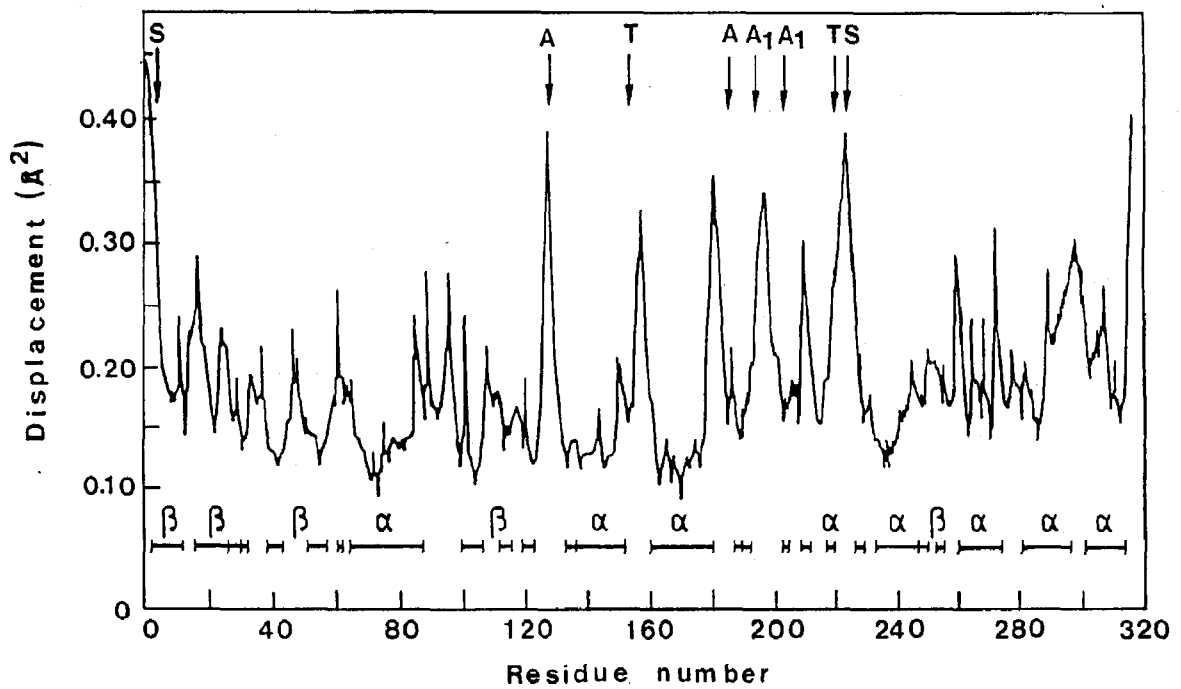

Fig. 3. Plot of the average main-chain temperature factors (solid line) along the polypeptide chain of thermolysin (adapted from refs. 69 and 102). Bars at the bottom of the figure indicate segments of secondary structure (helices and strands). Arrows denote sites of limited proteolysis or autolysis of thermolysin observed under different experimental conditions. Symbols (S, T, A and A_1) used to indicate sites of cleavage under different experimental conditions are those listed in the legend to fig. 2.

tal mobility, as given by the crystallographically determined thermal motion of the polypeptide chain of thermolysin. Fig. 3 shows the average mean-square displacement of C_α-carbons along the thermolysin polypeptide chain, providing a graphic image of the segmental mobility of the thermolysin molecule [69,110–112]. In this figure, locations of elements of regular secondary structure (α-helix and β-sheet) and of the peptide bonds cleaved by limited proteolysis are indicated. The most notable features of the graph in fig. 3 are that highest mobility occurs 'in the vicinity of residues 128, 180 and 225, as well as the amino and carboxyl termini' [69] and not in regions of regular secondary structure. The correlation between sites of highest mobility and those of limited proteolysis appears to be quite striking. All cleavage sites are observed not within segments of regular secondary structure (such as helices), but at loops or turns characterized by highest flexibility. The majority of cleavage sites occur within the most extended region of irregular conformation in the thermolysin molecule (region 180–230). Thus, taking the structure (fig. 1) and dynamic (fig. 3) features of the thermolysin molecule together, it can be concluded that the protein-protein recognition process between the protease and the globular protein substrate occurs at *exposed and flexible loops*. This observation is in keeping with the notion of a 'functional' role of loops in globular proteins, since loops are involved in several fundamental processes, such as phosphorylation, glycosylation, ion binding, antigenicity, hydroxylation, etc. [113]. In addition, introns in the eukaryotic genes of proteins are mainly located at boundaries between secondary structure elements, i.e., at loops or bends between domains and subdomains [114,115].

The results of these studies of limited proteolysis of thermolysin put on a firm basis the general notion [98] that limited proteolysis of a globular protein would be expected to occur at surface loops and random segments of polypeptide chains or at flexible hinges between protein domains rather than at internal loci or rigid elements of secondary structure such as helices or pleated sheets. The major result of the present study is that it emphasizes that segmental mobility is an essential part of the proteolytic event. One may envision the overall process of proteolysis in which the initial interaction of the globular protein with the protease involves recognition of a specific amino acid sequence of that site, after which some local conformational change takes place in order to make the idealized transition state of the cleavage reaction. Of note is the fact that the proteolytic process involves binding, at the active site of the protease, at primary and secondary sites, of peptide segments comprising, on average, six to eight amino acid residues [116,117].

7. Conclusions and perspectives

The reason for the continuing interest in thermophilic enzymes is related not only to the scientific curiosity of understanding the molecular mechanisms by which these macromolecules maintain their structure and function at high temperature, but also to the practical applications and developments connected with enzyme technology. Since the metabolic activities of thermophilic bacteria are similar to those of the mesophilic ones, it is expected that any enzyme already found in a mesophilic bacterium will most likely also be found in a thermophilic one. Thus, considering the wide variety of bacteria living in extreme environments at about 100 °C or above, it appears that there is an unlimited variety of stable enzymes that can be isolated and successfully used in biotechnology.

Genetic engineering techniques currently employed to facilitate production and usefulness of mesophilic enzymes, are also applicable to thermophilic enzymes. In fact, it has been demonstrated that genes from thermophiles can be expressed in mesophiles such as *E. coli*, resulting in the production of thermostable enzymes [44,46, 118–121], thus again demonstrating the intrinsic stability of thermophilic enzymes. It is also of interest to observe that, since the cellular proteins of *E. coli* are heat-denatured and precipitated, heat treatment of the *E. coli* cell extract containing the thermostable enzyme can provide a simple system for a partial and easy purification of the desired enzyme [119].

One major goal of current protein engineering studies is to improve enzyme properties such as resistance to heat, extremes of pH, organic solvents, proteolytic enzymes, etc. [122–124]. A limitation to the successful design of these new enzyme properties by site-directed mutagenesis is that the effects of amino acid replacements are not easy to predict and thus it is difficult to decide which amino acid substitutions should be made. In this respect, knowledge of structure-function-stability relationships in thermophilic enzymes would be of great help in gaining an understanding of the rules followed by nature in enhancing the enzyme stability and thus in devising strategies for improving mesophilic enzymes presently used in biotechnology.

Thermophilic microorganisms could be very important vehicles for the production of stable enzymes using genetic engineering methods. In fact, a method has been described [45] for rapidly generating thermostable enzyme variants by introducing the gene coding for a given mesophilic enzyme into a thermophile (*B. stearothermophilus*) and then selecting variants retaining the enzymatic activity at the higher growth temperatures of the thermophile. Using this procedure, thermostable variants of kanamycin nucleotidyltransferase have been successfully produced involving single amino acid replacements with respect to the wild-type enzyme [45].

Acknowledgments

This work was carried out with support of the Italian National Council of Research (CNR) and Commission of the European Communities, Biotechnology Action Program (contract no. BAP-0249-I). The author wishes to thank his co-workers who contributed to the various papers reviewed here and in particular Drs Claudio Vita and Daniele Dalzoppo. The excellent technical assistance of Marcello Zambonin and the expert typing of the manuscript by Mrs Elvira Piaia are also gratefully acknowledged.

References

1 R. Singleton and R.E. Amelunxen, Bacteriol. Rev. 37 (1973) 320.
2 H. Zuber, Enzymes and proteins from thermophilic microorganisms (Birkhäuser Verlag, Basel, 1976) Experientia Suppl. 26.
3 L.G. Ljungdahl and D. Sherod, in: Extreme environments: Mechanisms of microbial adaptation, ed. M.R. Heinrich (Academic Press, New York, 1976) p. 147.
4 R.E. Amelunxen and A.L. Murdock, in: Microbial life in extreme environments, ed. D.J. Kushner (Academic Press, New York, 1978) p. 217.
5 R.E. Amelunxen and A.L. Murdock, CRC Crit. Rev. Microbiol. 6 (1978) 343.
6 T.D. Brock, Thermophilic microorganisms and life at high temperatures (Springer Verlag, New York, 1978).
7 T.D. Brock, Science 230 (1985) 132.
8 F.M. Veronese, E. Boccù, O. Schiavon, C. Grandi and A. Fontana, J. Appl. Biochem. 6 (1984) 39.
9 A.M. Klibanov, Anal. Biochem. 93 (1979) 1.
10 B. Sonnleitner and A. Fiechter, Trends Biotechnol. 3 (1983) 74.
11 B. Sonnleitner, Adv. Biochem. Eng. Biotechnol. 28 (1984) 70.
12 D. Cowan, R. Daniel and H. Morgan, Trends Biotechnol. 3 (1985) 68.
13 B.S. Hartley and M.A. Payton, Biochem. Soc. Symp. 48 (1985) 133.
14 T.D. Brock and J.G. Zeikus, Thermophiles: General, molecular and applied microbiology (Wiley, New York, 1986).
15 J. Wiegel and L.B. Ljungdahl, CRC Crit. Biotechnol. 3 (1986) 39.
16 S. Endo, J. Ferment. Technol. 40 (1962) 346.
17 Y. Ohta, J. Biol. Chem. 242 (1967) 509.
18 K. Yutani, in: Enzymes and proteins from thermophilic microorganisms, ed. H. Zuber (Birkäuser Verlag, Basel, 1976) Experientia Suppl. 26, 91.
19 H. Drucker and S.L. Borchers, Arch. Biochem. Biophys. 147 (1971) 242.
20 R.S. Roche and G. Voordouw, CRC Crit. Rev. Biochem. 5 (1978) 1.
21 C. Grandi, C. Vita, D. Dalzoppo and A. Fontana, Int. J. Peptide Protein Res. 16 (1980) 327.
22 C. Tanford, Adv. Protein Chem. 23 (1968) 121.
23 C.N. Pace, CRC Crit. Rev. Biochem. 3 (1975) 1.
24 P.L. Privalov, Adv. Protein Chem. 33 (1979) 167.
25 S. Lapanje, Physicochemical aspects of protein denaturation (Wiley, New York, 1978).
26 G. Némethy and H.A. Scheraga, J. Phys. Chem. 66 (1962) 1773.
27 D.J. Merkler, G.K. Farrington and F.C. Wedler, Int. J. Peptide Protein Res. 18 (1981) 430.
28 A. Ikai, J. Biochem. (Tokyo) 88 (1980) 1895.
29 H.B. Bull and K. Breese, Arch. Biochem. Biophys. 158 (1973) 681.

30 R. Jaenicke, Annu. Rev. Biophys. Bioeng. 10 (1981) 1.
31 M.F. Perutz and H. Raidt, Nature 255 (1975) 256.
32 G. Biesecker, J.I. Harris, J.-C. Thierry, J.E. Walker and A.J. Wonacott, Nature 266 (1977) 328.
33 J.E. Walker, in Proteins: Structure, function and industrial applications, ed. E. Hofmann (Pergamon Press, Oxford, 1978) p. 211.
34 J.E. Walker, A.J. Wonacott and J.I. Harris, Eur. J. Biochem. 198 (1980) 581.
35 P. Argos, M.G. Rossman, U.M. Grau, M. Zuber, G. Frank and J.D. Tratschin, Biochemistry 18 (1979) 5698.
36 M.G. Grütter, R.B. Hawkes and B.W. Matthews, Nature 277 (1979) 667.
37 T. Alber, M.G. Grutter, T.M. Gray, J.A. Wozniak, L.H. Weaver, B.-L. Chen, E.N. Baker and B.W. Matthews, in: Protein structure, folding and design, ed. D.L. Oxender (A.R. Liss, New York, 1986) p. 307.
38 T. Alber and B.W. Matthews, Methods Enzymol. (1987) in the press.
39 T. Alber, S. Dao-Pin, J.A. Nye, D.C. Muchmore and B.W. Matthews, Biochemistry 26 (1987) 3754.
40 R. Hawkes, M.G. Grütter and J. Schellman, J. Mol. Biol. 175 (1984) 195.
41 J. Langridge, J. Bacteriol. 96 (1968) 1711.
42 K. Yutani, K. Ogasahara, Y. Sugino and A. Matsushiro, Nature 267 (1977) 274.
43 M. Matsumura, Y. Katakura, T. Imanaka and S. Aiba, J. Bacteriol. 160 (1984) 413.
44 M. Matsumura, S. Yasumura and S. Aiba, Nature 323 (1986) 356.
45 H. Liao, T. McKenzie and R. Hageman, Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 576.
46 T. Imanaka, M. Shibazaki and M. Takagi, Nature 324 (1986) 695.
47 F.R.N. Gurd and T.M. Rothgeb, Adv. Protein Chem. 33 (1979) 73.
48 J.A. McCammon and M. Karplus, Acc. Chem. Res. 16 (1983) 187.
49 D. Ringe and G.A. Petsko, Prog. Biophys. Mol. Biol. 45 (1985) 197.
50 A. Cooper, Sci. Prog. Oxford 66 (1981) 473.
51 R.M. Daniel, D.A. Cowan, H.W. Morgan and P. Curran, Biochem. J. 207 (1982) 641.
52 Y. Suzuki and T. Imai, Biochim. Biophys. Acta 705 (1982) 124.
53 M. Tsuboi, S. Ohta and M. Nakanishi, in: Biochemistry of thermophily, ed. S.M. Friedman (Academic Press, New York, 1978) p. 251.
54 G. Wagner and K. Wütrich, Nature 275 (1978) 247.
55 K. Wütrich, H. Roder and G. Wagner, in: Protein folding, ed. R. Jaenicke (Elsevier/North-Holland, Amsterdam, 1980) p. 549.
56 G. Careri, P. Fasella and E. Gratton, CRC Crit. Rev. Biochem. 3 (1975) 141.
57 R.J.P. Williams, Biol. Rev. 54 (1979) 389.
58 H. Zuber, in: Structural and functional aspects of enzyme catalysis, eds. H. Eggerer and R. Huber (Springer Verlag, Heidelberg, 1981) p. 114.
59 E. Stellwagen and L.D. Barnes, in: Enzymes and proteins from thermophilic microorganisms, ed. H. Zuber (Birkhäuser Verlag, Basel, 1976) Experienta Suppl. 26, 223.
60 K. Titani, M.A. Hermodson, L.H. Ericsson, K.A. Walsh and H. Neurath, Biochemistry 11 (1972) 2427.
61 M. Matsubara and J. Feder, in: The enzymes, 3rd edn., ed. P.D. Boyer (Academic Press, New York, 1971) vol. 3, p. 721.
62 J. Feder, L.R. Garrett and B.S. Wildi, Biochemistry 10 (1971) 4552.
63 J. Feder, Biochem. Biophys. Res. Commun. 32 (1968) 326.
64 R.L. Heinrikson, Methods Enzymol. 47 (1977) 175.
65 B.W. Matthews, J.N. Jansonius, P.M. Colman, B.P. Schoenborn and D. Dupourque, Nature 238 (1972) 37.
66 P.M. Colman, J.N. Jansonius and B.W. Matthews, J. Mol. Biol. 70 (1972) 701.
67 B.W. Matthews, L.H. Weaver and W.R. Kester, J. Biol. Chem. 249 (1974) 8030.
68 L.H. Weaver, W.R. Kester, L.F. Ten Eyck and B.W. Matthews, in: Enzymes and proteins from thermophilic microorganisms, ed. H. Zuber (Birkäuser Verlag, Basel, 1976) Experientia Suppl. 26, 31.
69 M.A. Holmes and B.W. Matthews, J. Mol. Biol. 160 (1982) 623.
70 W.R. Kester and B.W. Matthews, Biochemistry 16 (1977) 2506.
71 L.H. Weaver, W.R. Kester and B.W. Matthews, J. Mol. Biol. 114 (1977) 119.
72 M.C. Bolognesi and B.W. Matthews, J. Biol. Chem. 254 (1979) 634.
73 M.A. Holmes and B.W. Matthews, Biochemistry 20 (1981) 6912.
74 D.B. Wetlaufer, Proc. Natl. Acad. Sci. U.S.A. 70 (1973) 697.
75 G. Schultz, Angew. Chem. Int. Ed. Engl. 16 (1977) 23.
76 M.G. Rossmann and P. Argos, Annu. Rev. Biochem. 50 (1981) 497.
77 J. Janin and S.J. Wodak, Prog. Biophys. Mol. Biol. 42 (1983) 21.
78 S.J. Wodak and J. Janin, Biochemistry 20 (1981) 6544.
79 D.B. Wetlaufer, Adv. Protein Chem. 34 (1981) 61.
80 C. Vita, A. Fontana, J.R. Seeman and I.M. Chaiken, Biochemistry 21 (1979) 2016.
81 C. Vita, A. Fontana and I.M. Chaiken, Biochemistry 21 (1982) 2016.
82 C. Vita and A. Fontana, Biochemistry 21 (1982) 5196.
83 A. Fontana, C. Vita and I.M. Chaiken, Biopolymers 22 (1983) 69.
84 C. Vita, D. Dalzoppo and A. Fontana, Int. J. Peptide Protein Res. 21 (1983) 49.
85 C. Vita, D. Dalzoppo, S. Patti and A. Fontana, Int. J. Peptide Protein Res. 24 (1984) 104.
86 N.J. Greenfield and G.D. Fasman, Biochemistry 8 (1969) 4108.
87 J.B. Siegel, W.E. Steinmetz and G.L. Long, Anal. Biochem. 104 (1980) 160.

88 C. Vita, A. Fontana and I.M. Chaiken, Eur. J. Biochem. 151 (1985) 191.
89 M.J. Crumpton, in: The antigens, vol. 2, ed. M. Sela (Academic Press, New York, 1974) p. 1.
90 P.E.E. Todd, I.J. East and S.J. Leach, Trends Biochem. Sci. 7 (1982) 212.
91 D. Dalzoppo, C. Vita and A. Fontana, J. Mol. Biol. 182 (1985) 331.
92 A. Fontana, C. Vita and D. Dalzoppo, J. Biosci. 8 (1985) 57.
93 G.M. Crippen, J. Mol. Biol. 126 (1978) 315.
94 A.A. Rashin, Biochemistry 23 (1984) 2021.
95 E. Mihalyi, Application of proteolytic enzymes to protein structure studies, vol. 1 (CRC Press, West Palm Beach, 1978).
96 K.U. Linderstrøm-Lang, Cold Spring Harbor Symp. Quant. Biol. 14 (1949) 117.
97 H. Neurath and K.A. Walsh, Proc. Natl. Acad. Sci. U.S.A. 73 (1976) 3825.
98 H. Neurath, in: Protein folding, ed. R. Jaenicke (Elsevier/North-Holland, Amsterdam, 1980) p. 501.
99 C. Vita, D. Dalzoppo and A. Fontana, Biochemistry 24 (1985) 1798.
100 A. Fontana, G. Fassina, C. Vita, D. Dalzoppo, M. Zamai and M. Zambonin, in: Zinc enzymes, eds. I. Bertini, C. Luchinat, W. Maret and M. Zeppezauer (Birkäuser Verlag, Basel, 1986) p. 225.
101 G. Fassina, C. Vita, D. Dalzoppo, M. Zamai, M. Zambonin and A. Fontana, Eur. J. Biochem. 156 (1986) 221.
102 A. Fontana, G. Fassina, C. Vita, D. Dalzoppo, M. Zamai and M. Zambonin, Biochemistry 25 (1986) 1847.
103 C. Vita, D. Dalzoppo and A. Fontana, in: Macromolecular biorecognition: Principles and methods, eds. I.M. Chaiken, E. Chiancone, A. Fontana and P. Neri (Humana Press, Clifton, NJ, 1987) in the press.
104 K.A. Walsh, Y. Burstein and M.K. Pangburn, Methods Enzymol. 34 (1974) 435.
105 M.K. Pangburn and K.A. Walsh, Biochemistry 14 (1975) 4050.
106 G. Voordouw and R.S. Roche, Biochemistry 13 (1974) 5017.
107 P.L. Levy, M.K. Pangburn, Y. Burnstein, L.H. Ericsson, H. Neurath and K.A. Walsh, Proc. Natl. Acad. Sci. U.S.A. 72 (1975) 4341.
108 M.K. Pangburn, P.L. Levy, K.A. Walsh and H. Neurath, in: Enzymes and proteins from thermophilic microorganisms, ed. H. Zuber (Birkäuser Verlag, Basel, 1976) Experientia Suppl. 26, 19.
109 J.I. Harris and P. Roos, Biochem. J. 71 (1959) 445.
110 H. Frauenfelder, G.A. Petsko and D. Tsernoglou, Nature 280 (1979) 588.
111 P.J. Artymiuk, C.C.F. Blake, D.E.P. Grace, S.J. Oatley, D.C. Phillips and M.J.E. Sternberg, Nature 280 (1979) 563.
112 M.J.E. Sternberg, D.E.P. Grace and D.C. Phillips, J. Mol. Biol. 130 (1979) 231.
113 G.D. Rose, L.M. Gierash and J.A. Smith, Adv. Protein Chem. 37 (1985) 1.
114 N. Lonberg and W. Gilbert, Cell 40 (1985) 81.
115 C.-I. Brändén, in: Computer graphics and molecular modeling, eds. R. Fletterick and M. Zoller (Cold Spring Harbor Laboratory, NY, 1986) p. 450.
116 M. Ottesen, Annu. Rev. Biochem. 36 (1967) 55.
117 A. Berger and I. Schechter, Phil. Trans. R. Soc. Lond. 257 (1970) 249.
118 K. Nagahari, T. Koshikawa and K. Sakaguchi, Gene 10 (1980) 137.
119 T. Tanaka, N. Kawano and T. Oshima, J. Biochem. 89 (1981) 677.
120 G. Joliff, P. Béguin, M. Juy, J. Millet, A. Ryter, R. Poljak and J.-P. Aubert, Biotechnology 4 (1986) 896.
121 D.R. Love and M.B. Streiff, Biotechnology 5 (1987) 384.
122 W.H. Rastetter, Trends Biotechnol. 1 (1983) 80.
123 K. Ulmer, Science 219 (1983) 666.
124 R.J. Leatherbarrow and A.R. Fersht, Protein Eng. 1 (1986) 7.