Thermolysin Analysed

from our Molecular Biology Correspondent

THE latest enzyme to be given the full treatment, of sequencing and X-ray structure determination, is thermolysin, a protease from a thermophilic bacterium, which hydrolyses peptide chains on the carboxy-side of hydrophobic residues. It is a metalloenzyme, and contains both zinc and calcium, though only one zinc atom is indispensable for activity. It bears no relation to the familiar serine or thiol proteases, but in terms of molecular weight (which is 35,000), composition, and in particular the involvement of zinc in the catalytic process, it bears at the very least a passing resemblance to the animal enzyme, carboxypeptidase.

The sequence, however, which has been determined by Neurath and his group (Titani *et al.*, *Nature New Biology*, **238**, 35 ; 1972), belies this similarity, for a comparative analysis reveals no homologies. The only noteworthy feature of the sequence is the considerable clustering of residues of like charge in different parts of the chain. The electron density map at 2.3 Å resolution is shown by Matthews *et al.* (*ibid.*, 41), and reveals a now rather familiar pattern for proteins in this size range, two distinct globular units being separated by a deep groove, within which is lodged the zinc atom ; its location was established by extraction with a chelatine agent and replacement by mercury. Next to the zinc is a hydrophobic hole, into which a substrate side chain might be expected to fit, and so determine the specificity of the enzyme. A length of α-helix, from which emerge two ligands for the zinc, runs along the inside of the groove, and another zinc ligand comes from another long helix, which traverses the molecule. The two globular elements of the structure are made up of the C-terminal and N-terminal parts of the chain, and the authors draw attention to the relative prevalence of apparently independently folded units of this size—a faint echo perhaps of Svedberg's theory, which at one time had wide currency, that all proteins were made up of subunits of about 17,000 molecular weight.

One of the most cogent questions concerns the ligation state of the active site zinc atom. The ligands have been identified on the basis of the position of those residues in the sequence which are seen to have side chains approaching the zinc (Matthews *et al.*, *ibid.*), and comprise two histidines and a glutamic acid, disposed in more or less tetrahedral manner. These, as enzymologists will instantly recall, are the very same ligands as surround the zinc atom in carboxypeptidase. Matthews *et al.* observe that such common active site features in otherwise disparate enzymes of related specificity have been encountered before, in particular between the animal serine proteases and the bacterial enzyme, subtilisin. The similarity of the thermolysin to the carboxypeptidase site does not apparently end here, for the cavity behind the zinc atom contains a glutamic acid residue, perhaps corresponding to a similar side chain in carboxypeptidase. Close by, there is also an aspartate arginine ion pair, the arginine, which may again have a counterpart in carboxypeptidase, being apparently hydrogen-bonded to a histidine.

It is not certain whether the essential tyrosine of carboxypeptidase is likewise represented, but there is a tyrosine in the cleft, which may be hydrogen-bonded to one of the zinc ligands. Just how fundamental the similarity between the two enzymes really is in terms of mechanism is uncertain, because one can only guess at how the substrate fits into its slot. Certainly the overall geometry round the zinc, in spite of the identity of the ligand side chains, is distinctly different. From the reported presence of subsites in the enzyme, which recognize parts of the substrate on either side of the susceptible bond, it must be supposed that the cleft is lined with residues that participate in such interactions.

There remain the calcium binding sites: calcium ions are not, it seems, essential for activity, but reportedly stabilize the enzyme towards the high temperatures, at which the bacterium contrives to flourish. The X-ray structure reveals four calcium ions, two of them close together in a complex site, involving five side chain carboxylate groups, partly buried below the surface. The coordination of the calcium ions is thought to be octahedral. A third calcium ion is bound at the surface, apparently by way of two aspartates, and a fourth has also been identified at another aspartate side chain.

Insulin : in referring recently (*Nature*, **237**, 135 ; 1972) to work by Zimmerman, Kells and Yip on guinea-pig insulin, I allowed an error to creep in. The statement that, unlike bovine insulin, this protein does not associate "beyond the dimer" should of course read "to the dimer".

DNA Base Composition and Electrophoretic Mobility

IN *Nature New Biology* next Wednesday (July 19), Zeiger, Salomon, Wesley, Dingman and Peacock report a series of experiments of general significance to the many biochemists who separate DNAs by electrophoresis in polyacrylamide gels. Zeiger and his colleagues have obtained convincing evidence that the electrophoretic mobility of double stranded DNA is dependent not only on the size of the molecule but also on its base composition.

The first hint that this might be the case came when the group found that they could efficiently separate mouse satellite DNA from mouse main band DNA by polyacrylamide gel electrophoresis. The traditional way of achieving this separation is, of course, to centrifuge sheared mouse DNA and rely on differences in buoyant density, arising from differences in base composition, to effect the separation of A–T rich satellite DNA from G–C rich main band DNA.

Pursuing this lead Zeiger and his colleagues successfully attempted to separate by gel electrophoresis crab satellite DNA, which is almost pure poly dA–dT, from the main band DNA. Both mouse and crab satellite DNAs migrate more slowly than the main band DNA and their separation from the main band DNA must in part result from differences in base composition because molecules in the two separated bands show only small differences in molecular weight. These differences are not great enough to account for the electrophoretic separation.

Experiments with sheared and unsheared DNA from bacteria confirm this conclusion; for example, DNA extracted without shearing from three species of bacteria, *Bacillus subtilis* (42 per cent G+C), *Aerobacter aerogenes* (56 per cent G+C), and *Pseudomonas fluorescens* (63 per cent G+C), have electrophoretic mobilities respectively of 3.32, 3.53 and 3.66.