Sighting the cellular shredder

All biological systems produce waste. For every single cell of every single organism waste management is essential for survival. Robert Huber, Wolfgang Baumeister and colleagues have now determined the structure and probed the activity of the main cellular waste disposal unit for damaged or dangerous proteins, the 20S proteosome[1,2].

There are two major routes for proteins destined for the 'trash can' in eukaryotes. Those that enter the cell from the surrounding environment are degraded in a subcellular, membrane-bounded compartment called the lysosome (the interior of which is topologically 'outside' the cell). For cellular proteins, the route of degradation is, for the most part, through the ubiquitin-dependent proteolytic pathway, at the heart of which is the 26S ($\sim 2 \times 10^6$ M_r) proteosome complex—so named because of its sedimentation coefficient, 26 Svedbergs.

Many cellular proteins are marked for destruction by the addition of a molecular 'tag' consisting of a small protein known as ubiquitin (found in all organisms studied, hence the name). The tagged proteins are fed to the 26S proteosome complex where they are stripped of the ubiquitin tag (which is recycled) and shredded into small peptide fragments ranging from six to nine amino-acid residues in length. This is a surprisingly well-defined size range given the general lack of specificity of the 26S proteosome complex for its protein substrate. The peptide fragments released from the proteosome can then be degraded to their constituent amino acids by cytoplasmic proteinases.

The 26S proteosome complex is, by its very nature, a vital component in biological systems where the generation of short peptides or removal of potent regulatory proteins is important. For example, two β-subunits, LMP1 and LMP2, are encoded in the mammalian MHC class II locus. Their presence in the proteosome modifies its catalytic activity so that it cleaves proteins preferentially at hydrophobic and basic residues, generating peptides that can be readily presented on MHC class I molecules.

Biochemical studies and electron microscopy have revealed the gross anatomy of the eukaryotic 26S complex; it can be subdivided into a 20S ($\sim 7 \times 10^5$ M_r) proteosome 'catalytic core', where the business of protein degradation takes place, and a 19S complex, which is involved in regulating the activity of the core. The 20S core is a cylinder consisting of four seven-membered rings stacked on top of each other. Two outer rings of α-subunits bracket two inner rings of β-subunits, the resulting stoichiometry being $\alpha_7\beta_7\beta_7\alpha_7$. The different α- and β-subunits are the products of distinct but related genes. The 19S complexes cap both of the ends of the cylindrical 20S proteosome—giving the 26S complex a dumb-bell shaped appearance in the electron microscope. They are thought to bind and remove the ubiquitin moiety of the tagged protein, generate energy from ATP as part of the regulatory/degradation process and 'gate' the entrance to the 20S proteosome, unfolding the doomed proteins and threading them into the 20S catalytic core for destruction.

As well as being found in all higher organisms, an equivalent of the 20S proteosome core has been found in the thermophilic archaebacterium *Thermoplasma acidophilum*. Although it has the same overall structure as its eukaryotic counterpart, the prokaryote proteosome is much more simple: it does not have an associated 19S regulatory complex, and the α-subunits are all the product of one gene, as are all the β-subunits. This simplicity, as well as the additional stability of thermophilic proteins, has allowed Huber, Baumeister and colleagues to obtain crystals of the *Thermoplasma* proteosome and determine structure to a resolution of 3.4 Å, in the presence and absence of a potent peptide inhibitor, as well as to carry out a detailed biochemical analysis of its catalytic function.

The structure of the individual proteins is noteworthy (they possess a new fold and the

β-subunits have a novel active site) as is, of course, the macromolecular structure they form. The channel formed by the four rings, which runs the length of the cylinder, consists of three chambers; the outer two are constructed from the junctions between the α- and β-rings and the central chamber by the two β-rings. The cylindrical walls of the proteosome, unlike those of the superficially similar GroEL complex[3], are impenetrable from the outside. Thus access to both the outer and inner chambers is restricted solely by the small diameter of the entrance to outer α-rings (13 Å) and the entrance to inner β-rings (22 Å). Clearly a fully folded, globular protein (with an average diameter of 70 Å for a molecular weight of 50, 000 M_r) could not fit through such a small entrance. Rather, there is evidence to suggest that the polypeptide chain must be unfolded and fully extended to be threaded into the proteosome[4]. The entrance to the 20S proteosome, the outer chamber, the entrance to the inner chamber, and the inner chamber itself are all lined with hydrophobic residues which may function to maintain the fated polypeptide in an unfolded state. In eukaryotes the 19S complex seems to play an essential role in facilitating and regulating this threading. What the equivalent is in *Thermoplasma* is, as yet, unknown.

Guts of the machine

The catalytic heart of the proteosome is situated deep in the inner chamber. This much had already been surmised, yet the detailed nature of the catalytic machinery has remained something of a mystery. The α- and β-subunits are not homologous with any of the known families of proteinases and experiments have, until now, failed to identify the nature of the active sites. The structure of the complex between proteosome and peptide inhibitor reveals inhibitor molecules bound adjacent to the amino terminus of each of the 14 β-subunits lining the inner chamber. Each inhibitor is in close proximity to the side chain of the highly conserved Thr 1 residue. The site-directed mutagenesis studies also identify this Thr residue as critical for catalysis and suggest that it is the nucleophile that initiates cleavage of the peptide bond. Both the structural and mutagenesis studies hint that the base in the reaction (required to enhance the nucleophilic character of the attacking Thr hydroxyl group) is the α-amino group of the Thr 1 side chain itself: indeed, a similar catalytic mechanism has recently been identified for penicillin acylase[5].

The nature and disposition of the active sites in the inner chamber also provide an explanation for the ability of the proteosome to generate short peptide fragments which, in eukaryotes, are so vital for the MHC class I-mediated activation of the T-cell immune response. First, it seems that the substrate-binding pockets that flank the active site are relatively large and therefore non-specific, possibly allowing the substrate relative freedom of movement within the large inner chamber. Second, the distance between two neighbouring active sites is 28 Å, roughly the length of an extended seven–eight residue peptide. If the enzyme-bound intermediate in the reaction has a relatively long half-life, there is a possibility that the tethered and unfolded polypeptide chain will, as it continues to enter the inner chamber, encounter such a neighbouring active site and the second cleavage would then generate the observed peptide products.

The similarity of the *Thermoplamsa* 20S proteosome to the eukaryotic 20S proteosome indicates that the structure and biochemistry of the former will illuminate details of the latter. Clearly, the eukaryotic enzyme has acquired a whole raft of functions that mirror the much more complex biochemical environment in which it operates. Probing the sophistications of the 26S proteosome complex will inform us both about its function and the systems with which it interacts.

1. Lowe, J. *et al. Science* **268**, 533–539 (1995).
2. Seemuller, E. *et al. Science* **268**, 579–582 (1995).
3. Weissman, J.S., Sigler, P.B. & Horwich, A.L. *Science* **268**, 523–524 (1995).
4. Wenzel, T. & Baumeister, W. *Nature struct. Biol.* **2**, 199–204 (1995).
5. Duggleby, H.J. *et al. Nature* 373, 264–268 (1995).