Molecular characterization of the Drosophila 26S proteasome: subunit composition, subunit rearrangements in the regulatory complex and identification of O-glycosylated subunits.

Ph.D. Thesis

Máté Sümegi

Supervisor: Andor Udvardy Dr.

Szeged, 2003

Introduction

The homeostatic equilibrium in the eukaryotic cells is achieved by permanent destruction and rebuilding of the cellular components. The regulation of these processes is essential for the response to the genetical or environmental changes. Besides the synthesis and the destruction of highly abundant household proteins, the quantitative control of the less abundant but determinative regulatory proteins are very important for the process of the adaptation, development and reproduction. The controlled proteolysis of these regulatory proteins is one of the essential functions of the 26S proteasome. In addition, elimination of misfolded or damaged proteins arising as a consequence of different stresses is an other and very important function of the 26S proteasome.

The highly specific, controlled protein degradation is carried out by the ubiquitin – 26S proteasome pathway. The ubiquitin ligase system, belonging to this pathway, selects and labels the misfolded or unnecessary proteins by attaching ubiquitin monomers and forming a polyubiquitin polypeptide chain. The labelled proteins are degraded by a multiprotein complex, the 26S proteasome, to small oligopeptides and recyclable ubiquitin monomers.

Since the 26S proteasome is solely responsible for the intracellular proteolysis of the large majority of cellular proteins, it is not surprising that it is involved in certain steps of almost all cellular processes. It is involved in the degradation of misfolded proteins, in the quality control of the endoplasmic reticulum (ERAD), in the antigen presentation, in the cell-cycle and cell division control, in the maturing and degradation of transcription factors, oncogenes and tumour suppressors, and in many other processes.

The 26S proteasome exists in all eukaryotes and archea. It has been investigated in many species, but for practical reasons, most extensively in yeast, human and in Drosophila melanogaster. Besides the direct advantages of investigating proteasomes from human source, it has its obvious limits. The yeast as a research object is very important in the progress of this field, but unfortunately, it is not a good source of the complex. The research object used by our laboratory, the 0-12 hours old embryo of the Drosophila melanogaster, is a very rich source of the 26S proteasome. In addition, during the last hundred years Drosophila genetics produced a huge amount of information and methodical background facilitating the experimental use of this tiny creature.

To make the Drosophila system a widely accepted test object, detailed molecular characterization of the subunit composition of the Drosophila 26S proteasome was absolutely essential. This characterization is described in the first part of the thesis.

After the identification of all the subunits of the 26S proteasome, we investigated how the topology of the subunits changes after the reassembly of the 26S proteasome from its two main subcomplexes. To follow these changes, we exploited our monoclonal and polyclonal antibody collection. In the second part of our work, we determined the subunit specificity of all of these antibodies, and analysed the structural changes accompanying the assembly of the proteasome. During this work, we had an interesting observation leading us to suppose that some of the subunits of the 26S proteasome are posttranslationaly modified.

In the third set of our experiments we demonstrated the presence of oxygen-linked-β-N-acetyl-glucosamine modification on several subunits of the 26S proteasome.

The structure and enzymatic functions of the subcomplexes of the 26S proteasome

Regarding its structure and possible evolution, the 26S proteasome can be divided into two subcomplexes. One of its components is the 20S proteasome. This tube-like structure arises by stacking of four protein rings, each composed of seven similar α- or β-type subunits. The two central β rings are flanked by one α ring on each side. A ring can be composed of seven similar, or up to seven different α or β subunits in different species. Therefore the structural formula describing the composition varies between $\alpha_7\beta_7\beta_7\alpha_7$ and $\alpha_{1\text{-}7}\beta_{1\text{-}7}\beta_{1\text{-}7}\alpha_{1\text{-}7}$ from archea to eukaryota. Because of the simple symmetry and the easy of crystallization, good quality high-resolution X-ray diffraction picture has been published for the 20S proteasome. In the centre of this tube like particle there is a 2 nm wide central channel, on which there are three cave like flares at the boundaries of the protein rings. In the central nanocompartment between the β rings, active centres for three different peptidase activity are located, ensuring the degradation of peptide bonds following acidic, basic, and hydrophobic amino acid residues. The obvious advantage of closing the active centres of the peptidase activities inside the particle is to protect the cellular proteins not intended for degradation. This compartmentalisation is assisted by the channel gating mechanism of one of the α subunits. The protection of the cellular proteins is further enhanced by the fact that the central channel is too narrow to let the proteins get through without preliminary unfolding. The possibility of the spontaneous unfolding into a thread-like primer structure is very unlikely.

The selection of the ubiquitinated proteins and their feeding into the central channel of the 20S proteasome is the most important function of the 19S regulator complex. This subcomplex is built up from 19 different subunits. Unfortunately, it does not have any well-defined symmetry and cannot be crystallised, therefore X-ray crystallography failed to illuminate the fine structure of the subcomplex. Only computer-assisted electron microscopy provided a low resolution 2D picture of the particle. The structural studies were highly stimulated by the discovery of a yeast null mutant, in which the 19S regulator complex can be disintegrated into two different subcomplexes. One of these subcomplexes consists of a ring-like structure composed of six ATPase and three non-ATPase subunits. The ring is attached to the base of the 20S proteasome. One of the proposed roles of the ATPase ring is to open the

central channel of the 20S proteasome gated by one of the α subunits. Furthermore, its reverse chaperon activity ensures the unfolding of the substrate proteins.

The remaining subunits of the 19S regulatory complex constitute the lid subcomplex. The lid subcomplex is probably involved in the binding of the ubiquitinated substrate proteins and removing the ubiquitin chain from them. Different subunits of the lid may be responsible for the recognition and binding of certain subsets of the substrate proteins. The exact role of these subunits remains to be elucidated.

Materials and methods

Following the homogenisation of the 0-12 hour old embryos, we subjected the extract into a multistage centrifugation and ultra centrifugation procedure. Using the centrifugal separations interrupted with salt precipitation, we managed to eliminate the debris from the feeding plates, the unhomogenisated eggs, the nuclei of the cells, the mitochondria, the ribosomes, the nucleic acids, the membrane debris and the precipitated proteins arises during the freezing and storing.

This preliminary rough purification was followed by a multistep chromatographic fractionations. The large mass of the yolk proteins was removed on a hydroxilapatite column. This step was followed by a Fractogel EMD-DEAE (M) anion exchange and a Fractogel EMD-Heparin affinity chromatography. The elution position of the 26S proteasome, 19S regulatory complex and the 20S proteasome was tracked by immuno-dot blot procedure. In the final step, a gel filtration chromatography on Superose 6 Prep Grade column yielded an almost homogeneous proteasome fraction. The last two chromatographic steps were carried out in a computer controlled and highly reproducible FPLC chromatographic system.

Following the purification, we resolved the subunits on a 16-BAC/SDS 2D gel electrophoresis system, which has a loading puffer with better sample dissolving capability than that of the conventional IEF/SDS 2D electrophoresis method.

For mass spectrometric identification, the subunits were cut out from the gel, and after in-gel tryptic digestion our collaborators analysed the samples. The final identification of the subunit was accomplished by BLAST search in the NCBI protein database.

The subunit specific antibodies were assigned to the subunits separated on 16-BAC/SDS 2D gels followed by western blotting. Similar blotting technique was used to analyze the O-glycosylation of the subunits.

The integrity of the purified 26S proteasome, 19S regulatory complex and the 20S core particulum was checked by native polyacrylamide gel electrophoresis.

The proteolytic activity of the different components was visualised by fluorescent overlay assay.

The changes in the subunit rearrangements after the assembly of the 26S proteasome from its components were demonstrated by covalently cross-linking the neighbouring subunits of the assembled or unassembled proteasome with a bifunctional cross-linker. The

crosslinkig products were identified with 1D SDS denaturating gel electrophoresis followed by western blot experiments.

The electron microscopic localization of the p37A subunit was carried out with nanogold labelling technique in the laboratory of our collaborators. The enzymatic characterisation of the recombinant form of this subunit was made by fluorescent assays.

Results and discussion

Understanding the structure and the function of the 19S regulatory complex advances at a relatively slow pace. Initial studies identified 12 subunits in the *Drosophila* regulatory complex (RC) on 1D gels that were named p110 to p37. However, complete primary structures have been reported only for the p54 and p42C subunits.

Our first aim was to create the complete subunit catalogue of the 26S proteasome. This was achieved by separating the subunits of the purified 26S proteasomes by 2D gel electrophoresis and subsequent amino acid sequence analysis. In total, 19 subunits were found to constitute a single RC. Seventeen of them have homologues among the known yeast and mammalian RC subunits; hence, it can be assumed that none of these proteins is a contaminant and all are integral parts of the *Drosophila* RC. p37A, the only subunit missing in yeast, has homologues in mammalian 26S proteasomes. The mammalian subunits, S5b and p28, as well as yeast Rpn4, have not been consistently found in 26S proteasomes from other organisms. In the *Drosophila* RC, we found no homologues of these three subunits, in agreement with their absence in *Drosophila* EST cDNA data bases.

Most, but not all known RC subunits have been detected in all eukaryotic organisms investigated so far; hence, there seems to be a set of constitutive proteins that are essential for a functional 26S proteasome, and in addition, there should be facultative subunits. The latter may only be expressed in certain organisms or in a tissue-dependent manner, or only at specific developmental stages. Some of these facultative subunits may associate only transiently with the RC, and it will depend critically on the time point when a sample is taken whether they are detected or not. In fact, transient binding to the RC has been reported for Doa4 and Ap-uch, and also the bona fide subunit p54 is present in free and RC-bound form. Therefore, it appears unlikely that a universal number of all the subunits building the RC can be given.

Upon binding of ubiquitinated protein by the 26S proteasome, ubiquitin is usually recycled by means of deubiquitylating enzymes. We found the homologue of human deubiquitylating subunit UCH37 in *Drosophila* 26S proteasomes, which we named p37A, and

expressed in *E. coli* for further characterization. Like other recombinant proteins of this family, it cleaves the model substrate Ub-AMC and is inhibited by Ub-Al.

In addition, substrate preferences may differ between the free subunit and the subunit integrated into the RC. Recombinant p37A removes an intact ubiquitin chain from an ubiquitin-protein conjugate, whereas after embedding in the 19S complex it shortens an ubiquitin chain from the distal end by removing ubiquitin moieties one by one.

Since we found similar enzymatic activity and inhibition profiles with native 26S proteasomes and with recombinant p37A, we assume that p37A is at least in part responsible for the deubiquitylation of proteasome-bound conjugates. The fact that no homologue exists in the yeast genome suggests that p37A is not an essential subunit of the RC, and it cannot be excluded that there are additional subunits that exhibit deubiquitylating activity.

Having identified p37A as a bona fide component of the *Drosophila* RC, we have mapped its location by electron microscopy. To this end, we have taken advantage of the specific binding of the UCH inhibitor Ub-Al to its target. By coupling Ub-Al to 3-nm colloidal gold particles, a strong signal was generated. On averaged images, the Ub-Al gold conjugates map to the neck region of the dragon-head motif, i.e., the hinge between the base and the lid. This is the region where we also assume that p54 (Rpn10) is located. Thus, it appears that both the binding of multiubiquitin chains and their deubiquitylation are spatially closely related and perhaps also functionally coupled. One could envisage a scenario in which ubiquitylated proteins are first bound to the lid subcomplex. While being transferred to the base where the substrate is prepared for feeding into the 20Sproteasome, p54 prevents its escape, while p37A recycles the bound ubiquitin.

To study the topology of subunits within the RC of the *Drosophila* 26S proteasome, we identified the subunits recognized by the different subunit-specific monoclonal (or polyclonal) antibodies. The majority of the RC subunits are represented in our antibody library.

Cross-linking studies and electron microscopic observations revealed that the ATP dependent assembly of the 26S proteasome is accompanied by gross structural rearrangement of the RC. The ATPase subunits that have direct contacts with the 20S proteasome, are most extensively affected. The cross-linking pattern and its change following the assembly strongly suggest that three ATPase subunits (p48A, p42D and p42C) have multiple contacts with each other. For the assembled 26S proteasome, there is a broad cross-linked band, which reacts

with all three subunit-specific mAbs, indicating that these subunits are very closely spaced and can be cross-linked with each other efficiently. Disassembly of the 26S proteasome results in the displacement of these subunits, and the increased spatial distance between these subunits greatly reduces the efficiency of cross-linking at this contact point. Immunoprecipitation experiments confirmed the close proximity of these subunits. Protein overlay assays, as well as biochemical and genetic experiments published in the literature confirmed our results: strong interactions were demonstrated between subunits Rpt3/S6b/p48A–Rpt6/S8/p42C and Rpt4/S10bp/p42C–Rpt6/S8/p42D. Our cross-linking approach allowed demonstration not only of the interaction of these subunits, but also of the dynamic changes in these subunit interactions in the course of the assembly of the 26S proteasome.

The conformational changes accompanying the assembly of the 26S proteasome are not confined to the base of the regulatory complex. A very prominent change was observed in the regulatory lid subunit p42A. The appearance of the new band may be due to a direct contact of this subunit with the 20S proteasome, in which case the assumption that only the base subcomplex has direct contacts with the catalytic core is not valid. It is more probable, however, that the conformational rearrangement of the ATPase subunits generated this new contact. This suggests the neighbouring positions of p42A and the hexameric ATPase ring.

To support these biochemical data, electron microscopy and image analysis were carried out. For the first time 2D structures of the individual 19S regulatory complex are shown in this study, with resolutions better than 25Å according to the Fourier ring criterion. Hitherto, only rather indistinct and featureless structures of isolated and purified 19S complexes have been obtained. The *in situ* disassembly (following ATP depletion) of the 26S holoenzyme and imaging in unsupported ammonium molybdate-ice films allowed us to capture a number of intermediates following the detachment from the 20S core. Rapid ATP removal and the concomitant cryofixation of the samples minimize the risk of artefacts. Although the interpretation of these 2D structural data is hampered by the difficulty in discriminating between different orientations with respect to the viewing axis, the observed structural variants probably provide the morphological representation of the remodelling of the regulatory complex following the disassembly of the 26S proteasome.

The reconstitution of the 26 S proteasomes from highly purified RC and 20S proteasomes is inefficient because assembly factors are removed during the purification. In a

partially purified fraction of a *Drosophila* embryonic extract, however, the disassembly and the assembly of the 26 S proteasome are fully reversible. ATP depletion induces the disassembly of the complex, which can be fully reversed by the addition of an excess of ATP. Cross-linking studies of such a partially purified extract revealed that the structural rearrangements described above are not consequences of an artificial structural deterioration of the regulatory complex during the purification procedure. The changes are fully reversible and follow the assembly state of the proteasome, representing the remodelling process required for the assembly and the proper functioning of the 26S proteasome.

Due to our very limited knowledge of the molecular details of the catalytic cycle of the 26S proteasome, the interpretation of the structural changes observed by chemical cross-linking is very difficult. The subtle conformational rearrangement in the α-ring of the 20S proteasome may be associated with the gating of the channel. We can only speculate as to the roles of these rearrangements in the case of the regulatory complex. The spectacular assembly-dependent increase in the extent of cross-linking of subunits p42C, p42D and p48A with each other may represent a compaction of the hexameric ATPase ring. This may be required to match the size and/or the structure of the hexameric ATPase ring to that of the heptameric a- ring of the 20S proteasome. The energy required for this compaction is provided by the hydrolysis of the ATP, because a non-hydrolysable ATP analogue cannot support the increase in the extent of cross- linking of these subunits. The presence of the catalytic core is indispensable for this presumed compaction. In the free regulatory complex, ATP cannot induce the increase in the extent of cross- linking of these ATPase subunits. The physical interaction of the catalytic core and the regulatory complex may provide the physical support for this compaction.

It was recently shown that ATP hydrolysis modulates the association of the 26S proteasome with a multitude of proteasome-interacting proteins. It is reasonable to suppose that for each interaction the regulatory complex must adopt an ideal conformation, which is determined and induced by the interacting partner. In this scenario, the plasticity of the regulatory complex is the prerequisite structural basis of the functional redundancy of the proteasome.

The recognition that assembly of the 26S proteasome is not simply a passive docking of two rigid subcomplexes, but a process accompanied by substantial restructuring of the 26S proteasome, is the most important message of our results. In the course of the assembly, the

interacting subcomplexes mutually rearrange their structures so as to create the optimal conformation required for the assembly and the proper functioning of the 26S proteasome.

Recently the subunit interactions in the *Caenorhabditis elegans* and the *Saccharomyces cerevisiae* 26S proteasome have been studied by the yeast two-hybrid technique. Several subunit interactions, undetected by previous biochemical and genetic approaches, have been revealed. The interaction map generated by the yeast two-hybrid technique is very detailed, but provides only a static picture of potential subunit interactions. The cross-linking approach, used in this study, allowed an insight into the dynamic changes of subunit interactions during the assembly of the 26S proteasome. These observations indicate that several different experimental approaches will be required to map all the subunit contacts in the proteasome, before the crystal structure of this particle is solved.

Indirect evidences obtained from our cross-linking experiments suggested that posttranslational modifications might also induce rearrangements in the subunit contacts. We could prove the presence of a specific, hitherto unknown posttranslational modification on several proteasomal subunits.

Immunological detection with two different O-linked GlcNAc-specific monoclonal antibodies, together with lectin-binding experiments with wheat germ agglutinin, unequivocally proved the presence of O-linked N-acetylglucosamin modification on five RC and on almost all the 20S proteasome subunits.

The specificity of the detection of O-linked GlcNAc modification of proteins by monoclonal antibody or wheat germ agglutinin-binding methods has been well documented inthe literature. Our results are further supported by the observation that O-linked GlcNAc modification of the rat proteasomal catalytic core subunit C2 was recently demonstrated by immuno-affinity purification followed by mass spectrometry.

Analysis of the subunits of highly purified 26S proteasomes separated by 2D gel electrophoresis allowed us to demonstrate the presence of this post-translational modification on a multitude of proteasomal subunits.

With regard to the close and mutually exclusive dynamic interplay between O-glycosylation and O-phosphorylation, it is important to emphasize that the RC ATPase subunits p48A and p48B (that exhibit very strong reactions with the monoclonal antibodies) and subunit p56 (that reacts strongly with wheat germ agglutinin) belong to those RC subunits

that have been shown to be phosphorylated. In the 20S proteasome, phosphorylation of the α2, α4 and α7 subunits has been documented. Although the role of these post-translational modifications in the regulation of proteasomal function(s) is still unresolved, their significance in the integration of regulatory circuits affecting proteasomal function(s) should be considered in the future.