The Saccharomyces cerevisiae Cks protein Cks1 has a COOH-terminal glutamine-rich sequence not present in other homologues. Cks proteins domain swap to form dimers but unique to Cks1 is the anti-parallel arrangement of protomers within the dimer. Despite the differences in Cks1 compared with other Cks proteins, we find the domain swapping properties are very similar. However, aggregation of Cks1 occurs by a route distinct from the other Cks proteins studied to date. Cks1 formed fibrillar aggregates at room temperature and neutral pH. During this process, Cks1 underwent proteolytic cleavage at a trypsin-like site into two fragments, the globular Cks domain and the glutamine-rich COOH terminus. At high protein concentrations, the rate of fibril formation was the same as the rate of proteolysis. The dominant species present within the fibrils was the glutamine-rich sequence. Consistent with this result, fibril formation was enhanced by addition of trypsin. Moreover, a truncated variant lacking the glutamine-rich sequence did not form fibrils under the same conditions. A lag phase at low protein concentrations indicates that fibril formation occurs through a nucleation and growth mechanism. The aggregates appear to resemble amyloid fibrils, in that they show the typical cross-β x-ray diffraction pattern. Moreover, infrared spectroscopy data indicate that the glutamine side chains are hydrogen-bonded along the axis of the fibril. Our results indicate that the proteolytic reaction is the crucial step initiating aggregation and demonstrate that Cks1 is a simple, tunable model system for exploring aggregation mechanisms associated with polyglutamine deposition diseases.

The Cks (Cyclin-dependent kinase subunit) family of proteins is essential for control of the cell cycle in eukaryotes (1). The proteins are known to regulate the activity of cyclin-dependent kinases, although their underlying mechanism of action is not fully understood (2). The Cks structure consists of a four-stranded β-sheet capped at one end by two or three short α-helices (3) (Fig. 1A). Cks proteins undergo domain-swatchedimerization via the reciprocal exchange of the COOH-terminal β-strand, both in solution (4, 5) and in crystals (6, 7), although the very high dissociation constants observed suggest that these proteins do not domain switch in vivo. Domain swapping in Cks proteins is mediated by a highly conserved sequence, the so-called “hinge loop” (8) that links the exchanged β-strand with the rest of the protein (Fig. 1A). This part of the protein is the only region where there are differences in conformation between the monomer and dimer forms. The Cks homologue from Saccharomyces cerevisiae, Cks1, is the largest member of the Cks family and has a unique sequence at the COOH terminus containing 16 consecutive glutamine residues followed by 17 residues rich in glutamine, proline, and serine (Fig. 1B). Also unique is an anti-parallel arrangement of monomer subunits in the domain-swapped dimer form of the protein and a disulfide link between the two subunits as a result of a cysteine residue in the hinge loop (3) (Fig. 1A).

The expansion of CAG trinucleotide repeats coding for polyglutamine (poly(Q)) leads to neurodegenerative disorders such as Huntington disease that are characterized by the formation of insoluble, fibrillar aggregates in the brain (reviewed in Refs. 9 and 10). A threshold number of 35–45 repeats, varying somewhat among the diseases, is needed for pathogenicity. This threshold number is in good agreement with the length dependence of the aggregation kinetics of poly(Q) sequences in vitro (11–13). Aggregates are thought to cause disease either by direct toxicity (14) or through abnormal interactions with other cellular proteins by interfering with the regulation of gene transcription (15, 16). For some amyloid diseases it has been proposed that oligomeric intermediates formed early in the aggregation process may be even more harmful than mature fibrils (17) and it is speculated that this could also apply to poly(Q)-related disorders (18). At least nine inherited diseases are caused by CAG repeat expansions, but the proteins involved have little in common, either structurally or functionally, other than the poly(Q) region. Difficulties in working with the full-length proteins, on the one hand, and the low solubility of poly(Q) model peptides, on the other, mean that many details of the poly(Q) aggregation mechanism remain to be resolved. In particular, atomic information on the poly(Q) fibril structure has yet to be obtained and the role played by the non-poly(Q) regions of the proteins in the aggregation process is not yet defined. Poly(Q) aggregates are β-sheet rich and exhibit many of the defining features of amyloid fibrils (11, 13, 19–21). The aggregation of poly(Q) sequences is a nucleation-dependent process (20, 22, 23) similar to that seen for amyloidogenic peptides and proteins such as the Aβ peptide associated with Alzheimer disease (24). In these respects, expanded CAG repeat diseases appear to conform to the unified model of protein misfolding disorders (25, 26).

The two distinct features of Cks1, its glutamine-rich COOH terminus...
and the anti-parallel orientation of the monomers within the domain-swapped dimer, prompted us to investigate Cks1 folding and aggregation and to compare it with that of other Cks proteins that we have previously analyzed. In particular, we were interested in determining whether aggregation occurs through the poly(Q) region or through the globular region (either by domain swapping or otherwise). We show here that, despite the sequence and structural differences from the other Cks proteins, the domain swapping properties of Cks1 are conserved within members of this family (27, 28); by contrast, aggregation of Cks1 occurs by a route distinct from that of the Schizosaccharomyces pombe Cks protein, Suc1. Suc1 forms domain-swapped aggregates at neutral pH and room temperature, which subsequently convert to fibrillar aggregates. For Cks1, we observe fibril formation within a few days at neutral pH and room temperature. These fibrils, however, consist solely of the COOH-terminal glutamine-rich sequence of Cks1, residues 112–150. Fibril formation is dependent on the proteolytic cleavage of the COOH terminus and its rate is greatly increased by the addition of a protease that cleaves the protein at a point in the sequence just before the glutamine-rich region. The results are consistent with a model for the origin of poly(Q) disorders and a subset of other protein misfolding diseases, in which the critical step is the production of a fibrilligenic sequence by proteolysis.

MATERIALS AND METHODS

High purity urea was obtained from Rose Chemicals Ltd. (United Kingdom). All other chemicals were obtained from Sigma unless stated otherwise.

Protein Expression and Purification—The coding sequence for Cks1 was amplified from a S. cerevisiae cDNA library by PCR with primers introducing 5′ BamHI and 3′ EcoRI restriction sites. The PCR product was cloned into a pRSET(A) vector (Novagen) modified for the expression of a His6 NH2-terminal fusion protein with a thrombin cleavage site. Correct clones were identified by DNA sequencing (Oswell, University of Southampton). Site-directed mutagenesis (mutants C90V, C90S, P93A, E94P, P95A, and T118, the truncated variant at residue 118) was performed using the QuikChange kit (Stratagene) and was confirmed by DNA sequencing.

C41(DE3) cells were transformed with the plasmid and a small number of colonies were picked from the plate into flasks containing 1 liter of the growth medium.
Fibril Formation in Cks1

2TY and ampicillin media. The cells were grown at 37 °C to A600 ~ 0.6 and then induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside overnight. Cells were harvested by centrifugation at 5000 × g at 4 °C for 15 min, and the pellet was resuspended in 50 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl and 1 mM DTT. As Cks1 from S. cerevisiae contains one cysteine residue, 1 mM DTT was added to all buffers to prevent the formation of intermolecular disulfide bonds. Cks1 was purified by nickel-nitrilotriacetic acid affinity chromatography followed by size-exclusion chromatography. Protein samples were pure as judged by SDS-PAGE and mass spectrometry. The monomer and domain-swapped dimer forms of Cks1 were found to be well separated by gel filtration and were collected separately, as described previously for Suc1 (5).

Determination of Dissociation Constants for Domain Swapping—Determination of dissociation constants for the domain-swapped dimer of Cks1 was carried out as described previously (5). Briefly, wild-type or mutant Cks1 was incubated at 40 °C at different protein concentrations until the fraction of dimer remained constant with time. Samples were quickly cooled on ice for 5 min and the proportions of monomer and dimer were determined by analytical gel filtration (Amersham Biosciences, S75 Sepharose) at room temperature in 50 mM Tris-HCl (pH 7.5), 300 mM, 1 mM EDTA, 1 mM DTT. The data reflect the monomer-dimer equilibrium at 40 °C as the samples do not re-equilibrate significantly at room temperature over the course of the experiment. A plot of the square of the monomer concentration versus the dimer concentration was fitted to a straight line, the slope of which gives the dissociation constant.

Transmission Electron Microscopy—Protein samples (at 1.8 mM concentration) were incubated for varying lengths of time at 25 °C in 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT. For electron microscopy, samples were diluted 100 times and applied to Formvar-coated nickel grids. The blotted sample was negatively stained with 2% (w/v) aqueous uranyl acetate solution and examined in a Philips CM100 electron microscope operated at 80 kV.

Experiments with Protease and Protease Inhibitors—1.7 mM Cks1 was incubated either with 0.1 mg/ml trypsin or a protease inhibitor mixture (1 Complete tablet/100 ml, Roche Diagnostics) at room temperature. The samples were then analyzed by mass spectrometry and electron microscopy after 1 and 2 days, respectively.

Mass Spectrometry—Mass spectra were collected using a quadrupole time-of-flight-2 mass spectrometer equipped with a nanoflow Z-spray source (Micromass, Manchester, UK) as described previously (29).

Amino Acid Analysis—High-order oligomers of larger than 8 S formed during aggregation of Cks1 were sedimented in a Beckman Optima TLX ultracentrifuge. The supernatant was carefully removed from each sample and the pellet resuspended in water. The two protein fractions were hydrolyzed and analyzed using the standard ion-exchange ninhydrin procedure (30).

X-ray Fiber Diffraction—Fibrils were purified by sedimentation (as described above) and resuspended at 5–10 mg/ml. X-ray specimens were aligned using a stretch frame by air-drying these samples between two wax-filled capillary ends. Fiber diffraction images were collected on a R-AXIS IV image plate x-ray detector mounted on a Rigaku RU200 rotating anode source of CuKα (1.5418 Å). Images were analyzed by the Java application “Crystal Clear” software.

Fourier Transform Infrared Spectroscopy—Samples were placed in a BioATRCell™ II equipped with a silicon crystal. The cell was equilibrated at 25 °C by using a water bath and infrared spectra were recorded with a Bruker Equinox 55 Fourier transform infrared spectrometer equipped with a liquid nitrogen-cooled mercury cadmium telluride detector. The sample compartment was continuously purged with dry air and 256 interferograms were co-added after registration with a resolution of 2 cm⁻¹. The buffer spectrum was recorded and subsequently subtracted from a protein sample using Protein Dynamics software (Bruker). Corrections for water vapor in the atmosphere were applied where necessary.

RESULTS

Domain Swapping Behavior of Cks1—To determine whether or not domain swapping leads to aggregation of Cks1, we first quantified the domain swapping propensity of the wild-type protein and of a number of mutants including a variant lacking the glutamine-rich COOH-terminal region. All measurements of domain swapping behavior were carried out in the presence of reducing agent so that the intrinsic propensity of the protein to domain swap could be investigated without the complicating factor of intermolecular disulfide bond formation. The interconversion between the monomeric form of Cks1 and its domain-swapped dimer was found to be very slow at room temperature (several hours, depending on the protein concentration), as observed previously for the S. pombe Cks protein Suc1 (5) and therefore the Kd was determined at 40 °C to enable the equilibrium distribution of the two forms of the protein to be reached on a reasonable time scale. Wild-type Cks1 is fully folded at this temperature as it is well below its midpoint of denaturation (58 °C from thermal unfolding monitored by CD and fluorescence, data not shown). The same situation holds for the mutants studied here (Table 1) as only small changes (1–2 °C) in the midpoint of denaturation were observed. The proportion of monomer and dimer in each sample was determined by analytical gel filtration, carried out at room temperature so that the distribution of the two species is not perturbed during the time required for these experiments. The monomer and dimer are well separated in the elution profile (Fig. 2A), and the Kd was calculated by measuring the monomer and dimer content of samples at different protein concentrations. A plot of the dimer concentration versus the square of the monomer concentration is linear, indicating that equilibrium was effectively achieved for all protein concentrations investigated (Fig. 2B), and the slope of the line gives the Kd values for the wild-type and mutant Cks1 dimer (Table 1).

To probe the domain swapping properties further, we mutated residues in the hinge loop (residues 91 to 99, sequence HAPEPHILL), the only region of the protein that adopts a different structure in the monomer compared with the dimer. We also investigated whether the glutamine-rich COOH-terminal region modulates the domain swapping propensity of Cks1. Dissociation constants of the domain-swapped dimer for wild-type and mutant Cks1 are given in Table 1 together with those previously measured for Suc1 for comparison. Substitution of the
The Poly(Q) Sequence of Cks1 Is the Origin of Fibril Formation—

The Cks1 monomer and dimer were incubated for several days at room temperature and under near physiological conditions (pH 7.5, 1 mM EDTA, 1 mM DTT) and a protein concentration of 1.8 mM. Samples from the solution were taken at regular intervals and examined using negative stain transmission electron microscopy (EM). The samples were found to form clear gels after several days of incubation, a phenomenon that is often indicative of aggregation. The EM images show that both monomer and dimer samples initially form amorphous aggregates, but long, unbranched fibrils appear after a few days (Fig. 3). The fibrils are typically 1–5 μm in length and 20 nm in width. The morphology appears very similar whether the fibrils are formed from the monomeric form or the dimeric form of Cks1 (Fig. 3). Neither the monomeric nor the dimeric form of the truncated variant of Cks1 lacking the poly(Q) region (T118) forms observable fibrils under the conditions used here, although amorphous aggregates were evident.

Proteolytic Release of the Poly(Q) Sequence Initiates Fibril Formation—

The oligomerization state of Cks1 during the fibril incubation period was probed by analytical gel filtration. The chromatogram obtained for monomeric wild-type Cks1 at the beginning of the incubation period (Fig. 4A) shows a single peak with an elution volume of 16 ml. After 4 days of incubation, two further peaks appear at 14.5 and 17.6 ml. The former peak corresponds to the domain-swapped dimer. The latter peak was subsequently shown by mass spectrometry to correspond to a truncated monomer (M*) (see below). Some of the dimer molecules are also truncated, as is evident from the additional peak at 15.4 ml (D*) observed in the sample that was incubated for 10 days. The elution volumes of the truncated monomer and dimer correspond closely to those of the monomer and dimer of the T118 variant (17.1 and 15.9 ml, respectively). Analysis by mass spectrometry (Fig. 6A) shows that cleavage of the protein takes place at residue Arg111, just prior to the poly(Q) region, where the most COOH-terminal trypsin site in the sequence is located (Fig. 1B). The fact that cleavage occurs in the samples may be attributed to trace quantities of proteases remaining after purification. No cleavage was observed for T118, even though the trypsin site at position Arg111 is still present in this truncated variant. Possible explanations for this protease resistance are given under “Discussion.” The mass spectrometry data also reveal that heterodimers composed of full-length and truncated protein molecules were also present in the solutions after 1 day, as were oligomers of the poly(Q) fragment produced by proteolysis. Fig. 4B shows a plot of the total areas of the elution peaks from the truncated monomeric and dimeric species as a function of incubation time. The time courses of the proteolytic cleavage of the wild-type and mutants of Cks1 (Fig. 4B) reflect those observed for gel formation; the wild-type and P93A monomer form gels after 7 days of incubation, the E94P monomer after 12 days and the E94P dimer after 15 days. The results are consistent with the fibrils being composed of the cleaved poly(Q) region, and they indicate that the rate-limiting step of fibril formation under the conditions used here is the proteolytic reaction through which this region is generated. Fibrils Formed in the Presence of Trypsin Consist Solely of the Poly(Q) Sequence—Fibril formation at a concentration of 1.7 mM Cks1 was monitored both in the presence of protease inhibitors and presence of

FIGURE 2. Determination of the dissociation constant for the domain-swapped dimer of Cks1. A, representative elution profile of wild-type Cks1 on an analytical Superdex 75 column in 50 mM Tris–HCl buffer (pH 7.5), 1 mM DTT, 1 mM EDTA, showing monomer and dimer peaks. B, plot of the square of the concentration of monomer versus the concentration of dimer for Cks1 PA95.

FIGURE 3. Fibril formation by the Cks1 monomer and domain-swapped dimer. Negative stain transmission electron micrographs of Cks1 fibrils generated from incubation of protein, at a concentration of 1.8 mM for monomeric wild-type Cks1 and the truncated variant T118 and 0.9 mM for the dimeric wild-type protein (corresponding to 1.8 mM monomers), at room temperature in 50 mM Tris–HCl buffer (pH 7.5), 1 mM DTT, 1 mM EDTA.
Fibril Formation in Cks1

Biophysical Characterization of Poly(Q) Fibrils—Fibrils produced by incubation of Cks1 with trypsin at pH 7.5 remained intact and in solution below pH 3.9 and further biophysical analysis of the fibrils was carried out at pH 2.0. The x-ray diffraction pattern of the fibrils, measured following their alignment by means of a stretch frame, at this pH contained two main reflections (Fig. 7A). One is a 4.7-Å reflection on the meridian, a reflection commonly observed for amyloid fibrils and attributed to the spacing between strands in a β-sheet structure where the strands are oriented perpendicular to the fibril axis (31). This reflection is dominant and sharp, as expected because this spacing is essentially sequence independent and there is an effectively infinite array of well oriented strands in each fibril (32).

As well as the 4.7-Å meridional reflection, classical amyloid fibrils have an equatorial reflection at 10–12 Å, corresponding to the distance between the arrays of β-sheets in the fibrils. The position of this reflection is dependent on the steric bulk of the component amino acid side chains (33), and it has been found to be weak in many samples of amyloid fibrils and absent in the diffraction patterns of poly(Q) and poly(N) fibers examined by Perutz et al. (34). The intensity of a reflection of this type depends on the number of β-sheets that are stacked together in each fibril (34) and it can also be weak because the fibrils are rotationally averaged around their long axis (32). In the diffraction pattern of the fibrils from the poly(Q)-rich COOH-terminal fragment of Cks1, this equatorial reflection, expected at 10–12 Å is absent. Instead, we observe a weak equatorial reflection at 8.3 Å, which corresponds to a similar reflection reported by Perutz et al. (34, 35).

Fig. 7B shows the CD spectrum of full-length Cks1 under native conditions at pH 7.5 (dotted line) and poly(Q) fibrils from Cks1 at pH 2.0 (solid line). The CD spectrum of the fibrils from Cks1 at pH 2.0 is typical of that observed for β-sheet structures, and has no apparent contributions from any other type of secondary structure (Fig. 7B). The same observation has been reported previously for a poly(N) peptide (20) and the poly(Q) and poly(N)-rich regions of the Sup35 protein of yeast prions (20). The CD data are supported by Fourier transform infrared spectroscopy measurements that show a band at ~1645 cm⁻¹ for full-length Cks1 at pH 7.5 (Fig. 7C, dotted line). By contrast, a band at ~1622 cm⁻¹, typical of β-sheet structure, is obtained for the poly(Q) fibrils from Cks1 at pH 7.5 and 2.0 (Fig. 7C, solid lines). The amide I band in the fibril spectra obtained at pH 2.0 is further characterized by two well resolved bands at ~1608 and ~1656 cm⁻¹, which can be assigned to the C = O stretching and NH₂ deformation vibrations of the glutamine side chains (36, 37). The positions of these bands are slightly shifted to lower and higher wave numbers, respectively, compared with the positions observed in spectra of free glutamine (36, 37). The shifts and high intensities of the absorption bands of the side chains indicate that the latter are involved in hydrogen bonding interactions. This conclusion is in agreement with previous findings by Krull et al. (38). Also in Fig. 7C is the amide I band at pH 7.5, which shows a higher absorbance at ~1640 cm⁻¹ than at pH 2.0 in addition to the three peaks observed at pH 2.0. The ~1640 cm⁻¹ peak is probably an absorption band of the peptide –NH₂ end groups and its increased intensity at higher pH values reflects the pH dependence of its extinction coefficient (36, 37). Furthermore, absorption in this region of the IR spectrum is generally attributed to the disordered structure, although it is possible that non-fibrillar species contribute to this absorbance and are removed by lowering the pH and further purification of the sample. The smaller bandwidth at pH 2.0 is suggestive of an increased homogeneity of the vibrations, which is likely to reflect the fact that the fibrils remain in solution only at low pH.
A similar spectrum to that reported here at pH 7.5 has been observed for fibrils of ataxin-3 that contain a 78-residue glutamine repeat region (39).

DISCUSSION

Cks1 and Other Cks Proteins Show Similar Domain Swapping Behavior—The orientation of the monomers in the crystal structure of the domain-swapped dimer of Cks1 is different from that observed for the dimers of the other Cks proteins (3, 6, 7, 40, 41) (Fig. 1). The Cks1 dimer is also unique in that it has a disulfide bond linking the two monomers. Despite these differences, mutations in the hinge loop alter the domain swapping propensity in similar ways in all three of the Cks proteins studied so far (5, 28). Mutation to alanine of the first proline residue in the hinge loop (position 93 in Cks1) increases the $K_d$ of the dimer formation by several orders of magnitude, whereas mutation to alanine of the second proline residue in the hinge loop (position 95 in Cks1) decreases the $K_d$ significantly; it has been shown previously that the first proline residue creates strain in the hinge loop conformation of the monomer, and that the second proline residue creates strain in the dimer (5). Substitution of the glutamate residue in the hinge loop (position 94 in Cks1) by a proline, resulting in a repeat of three proline residues, decreases the $K_d$ of the dimer by several orders of magnitude relative to the wild-type protein. This effect has been rationalized previously in terms of the additional strain likely to be generated in the hinge loop of the monomer by such a sequence of residues (5). By contrast, mutation of the cysteine residue at position 90 has only a small effect on the $K_d$, and the equivalent position in the other Cks proteins is occupied by a small hydrophobic residue, such as isoleucine or valine.

Interestingly, although the presence of the poly(Q) sequence does not alter significantly the free energy of unfolding of the protein (data not shown) it does decrease the domain swapping propensity by almost 3-fold. Poly(Q) sequences have been inserted in two proteins, chymotrypsin inhibitor 2 (42) and RNase A (43). Unlike Cks1, the insertions were in the middle of the proteins, in the hinge loop region, and the result was that domain swapping was promoted with poly(Q) sequences assembling from different molecules assembling together. In Cks proteins the swapped domain is the most COOH-terminal $eta$-strand and the poly(Q) sequence is located COOH-terminal to this strand in Cks1; consequently, the poly(Q) sequence should not be conformationally constrained in either monomer or dimer forms and so should not affect the relative stability of the two forms. Moreover, the monomers adopt an anti-parallel arrangement according to the crystal structure of the Cks1 dimer and so the two poly(Q) regions extend from the globular dimer in opposite directions and would not be expected to interact. It is therefore not clear why the poly(Q) sequence has an effect on the domain swapping propensity of Cks1.

FIGURE 5. EM data for Cks1 incubated in the presence of a protease or protease inhibitors. 1.7 mM Cks1 was incubated with a protease inhibitor mixture (A) or 0.1 mg/ml of trypsin (B) in 50 mM Tris-HCl buffer (pH 7.5), 1 mM EDTA, 1 mM DTT for 2 days at room temperature. The negative stain electron micrographs of the fibrils shown were produced from a 100-fold diluted solution of the incubated protein.

FIGURE 6. ESI-MS spectra and amino acid analysis of wild-type Cks1 after incubation at room temperature. A, Cks1 at a concentration of 1.7 mM was incubated for 2 days at pH 7.5. The spectrum is dominated by the charge series corresponding to full-length Cks1 (A series) and the cleaved Cks1 (B series) species. In addition, a heterodimer consisting of cleaved and uncleaved Cks1 and an oligomer consisting of 11 poly(Q)-rich COOH-terminal peptides of Cks1 can be detected (series C and D, respectively). B, Cks1 was incubated in the presence of 0.1 mg/ml of trypsin for 24 h. All fragments predicted for the fully digested Cks1 protein were identified, and are indicated, except those consisting of the COOH-terminal poly(Q) region; by contrast, peaks corresponding to the molecular mass of the COOH-terminal poly(Q) peptide fragment (residues 112–150) are absent, consistent with their incorporation into large aggregates. C, aggregates > 8 S from the trypsin-treated Cks1 solution as described in B were sedimented by ultracentrifugation. The compositions of the supernatant (gray columns) and pellet (black columns) were determined by amino acid analysis. The observed values fit well with the expected mole ratios (mole amino acid/mol protein) for each amino acid (expected values are indicated by black bars), assuming that the supernatant and the pellet consist of the digested fragments from the globular Cks1 domain (lacking the poly(Q) fragment) and the poly(Q)-rich region, respectively.
Fibril Formation in Cks1

The Poly(Q) Sequence of Cks1 Forms Fibrillar Aggregates following Proteolytic Cleavage—For Cks1, two mechanisms of self-association suggest themselves: the first involves the globular region and takes place via domain swapping and the second involves the glutamine-rich COOH-terminal region. The results in this study show that aggregation of the glutamine-rich residues can generate well ordered fibril structures at neutral pH and room temperature, whereas domain swapping in the globular region does not appear to generate ready formation of aggregates under the conditions explored in this study. At millimolar protein concentrations, proteolysis is the rate-limiting step in fibril formation and the rate of fibril formation can be greatly increased by the addition of trypsin. At submillimolar protein concentrations, a significant lag phase in fibril formation is observed, consistent with the nucleation mechanism that has been proposed for the formation of fibrils from poly(Q) sequences (13, 44), and indeed more generally for amyloid fibrils (24).

In a study of poly(Q) peptides containing between 15 and 44 glutamine residues flanked by 2 lysines on each side, both the nucleation lag time and the critical protein concentration required for aggregation were found to depend strongly on the glutamine repeat length (13). At 10 μM, the Gln15 peptide exhibited a lag time of more than 10 days, whereas aggregation of the Gln44 peptide occurred after only a few hours. Here we report that a 39-residue COOH-terminal fragment of Cks1, produced upon proteolytic cleavage and containing a stretch of 16 consecutive glutamine residues, does not form fibrils for at least 2 days when incubated at submillimolar concentrations. By contrast, at millimolar protein concentrations precipitation occurs within a few hours with well defined fibrils becoming visible after only 2 days. The poly(Q) stretch appears to remain in solution at millimolar protein concentrations, however, as long as it is attached to the globular domain in the context of the full-length Cks1. Proteolytic cleavage is therefore the critical first step in the overall aggregation process described here, as it results in a high local concentrations of a relatively short poly(Q) stretch that would otherwise have little tendency to aggregate. One question that arises from our results concerns the origin of the proteolytic activity that is observed to recognize the trypsin-like target site on Cks1 but is not affected by trypsin inhibitors (compare Fig. 3 showing EM of Cks1 alone with Fig. 5 showing EM of Cks1 incubated with inhibitor). Interestingly in this regard is the observation that the truncated variant T118 lacking the poly(Q) sequence is resistant to proteolysis although it still contains the cleavage site at position Arg 111. There are two possible explanations. The first is that Cks1 itself, in combination with a poly(Q)-rich COOH-terminal extension, has some intramolecular cleavage activity. This possibility is rather unlikely because Cks1 is not known to possess such activity against other proteins, nor is any other Cks protein. The second explanation is that a protease co-purifies only with the poly(Q)-rich Cks protein. This protease is thought to be Arg 111 and is shown to have some proteolytic activity against a few selected proteins, but it cannot be ruled out that this protease is not specific for Cks1.

The Structure of the Poly(Q) Fibrils—Fibrils of proteins containing an elongated polyglutamine sequence have been shown to bind Congo red and exhibit green birefringence when examined under polarized light (11, 39); these structures have been classified as amyloid fibrils on the basis of these properties. However, the molecular structure is unknown and several models have been proposed for the fibrillar structures formed by many different types of protein, including cross-β and β-helix structures. Perutz and co-workers (34) suggested the latter as a model for poly(Q) aggregates and its possible relevance to other amyloid fibrils has been discussed elsewhere (45, 46). The β-helix model is based on an analysis of the x-ray diffraction pattern of Asp2-Gln15-Lys2 in which the 10–12-Å equatorial reflection, characteristic of the cross-β structure (31), is absent but an additional equatorial reflection at 8.3 Å can be observed; the origin of this latter reflection is still a matter of debate (35, 47). In the present work we have obtained a diffraction pattern from fibrils derived from Cks1 that is similar to that of Asp2-Gln15-Lys2, and in addition, the infrared spectrum has extra features...
Fibril Formation in Cks1

compared with those of typical cross-β amyloid structures, such as found with insulin fibrils (48). Most amyloid fibrils contain mainly the β-sheet structure as indicated by a strong component in the Fourier transform infrared spectrum between 1638 and 1622 cm⁻¹ (49), without significant absorption that is typical of turns or loops. By contrast, in the case of Cks1 fibrils the β-sheet band at 1622 cm⁻¹ is accompanied by two almost equally intense bands at 1608 and 1656 cm⁻¹, which arise from the glutamine side chains. We have interpreted the intensity and position of the glutamine side chain bands in the amide I region of the IR spectrum in terms of hydrogen bonding. Two recent models for poly(Q) fibrils and the Q/N-rich Sup35 peptide (35, 50), based on x-ray diffraction data, offer an explanation of how these interactions might arise. In these models the glutamine and asparagine side chains are interdigitated allowing tight packing. Moreover, these side chains form interamidine hydrogen bonds along the axis of the fibril, and we suggest that these features could be the origin of the additional IR bands; they may also account for the 8.3 Å reflection and the higher order broad equa-
tional reflection at roughly 3.9 Å.

Cks1 as a Model System for Studying Mechanisms of Aggregation Associated with Disease—Naturally occurring poly(Q) sequences are found in a number of proteins, including several transcription factors. They are usually less than 25 residues in length, i.e. well below the human threshold of about 35 residues that is associated with human pathologies. Cks1 from S. cerevisiae is unique among the Cks family of proteins in having a glutamine-rich sequence at its COOH terminus. The remainder of the sequence, constituting the globular domain, is very highly conserved and it has been shown that the COOH-terminal region of Cks1 is not required for its normal function (3). It therefore seems unlikely that the glutamine-rich sequence plays any functional role. However, to date no studies have addressed whether proteolytic cleavage and/or aggregation of Cks1 occur in vivo.

Several studies indicate that proteolytic release of a fibrilligenic peptide is an important aspect of a number of protein misfolding diseases. Caspases and calpains, proteases important in apoptosis, have been shown to release the poly(Q)-containing fragment of the huntingtin protein and the Alzheimer Aβ peptide is released after proteolysis of the amyloid precursor protein by secretase enzymes (51). Also, it has been consistently observed that antibodies selective for the NH₂-terminal region of huntingtin, which includes the poly(Q) region, but not those that recognize the middle or COOH-terminal regions of the sequence, bind to the proteinaceous inclusions in Huntington disease. There are multiple ways in which a poly(Q) peptide fragment might convert into a pathogenic form; for example, as well as self-associating, the cleaved peptide might co-aggregate with the uncleaved protein or with other poly(Q)-containing proteins. This latter effect could perhaps explain the toxicity associated with the aggregation phenomenon, for example, by transcriptional dysregulation (16, 52), and indeed poly(Q)-containing transcription factors have been detected in neuronal inclusions (15, 53–55). Poly(Q) fragments may also interfere with other cellular processes such as ubiquitin-mediated protein degradation by the proteasome (56). It has recently been shown that the proteasome cannot cleave long sequences of poly(Q) residues (57), suggesting that the occa-
sional failure of partially degraded products to exit the proteasome could perhaps obstruct its central channel, causing malfunction. The proteasome may also contribute to the aggregation process by releasing poly(Q) fragments with a greater tendency to aggregate than the full-length poly(Q) proteins (57).

There are several problems inherent in characterizing the assembly mechanism of poly(Q) sequences: on the one hand, the peptides in isolation are intractable because they are insoluble even at low concen-

REFERENCES

1. Harper, J. W. (2001) Curr. Biol. 11, R431–R435
2. Pins, J. (1996) Curr. Biol. 6, 1399–1402
3. Bourne, Y., Watson, M. H., Arvai, A. S., Bernstein, S. L., Reed, S. I., and Tainer, J. A. (2000) Structure 8, 841–850
4. Birck, C., Raynaud-Messina, B., and Samama, J. P. (1995) FEBS Lett. 363, 145–150
5. Rousseau, F., Schymkowitz, J. W., Wilkinson, H. R., and Itzhaki, L. S. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 5596–5601
6. Parge, H. E., Arvai, A. S., Muratli, D. J., Reed, S. I., and Tainer, J. A. (1993) Science 262, 387–395
7. Endicott, J. A., Noble, M. E., Garman, E. F., Brown, N., Rasmussen, B., Nurse, P., and Johnson, L. N. (1995) EMBO J. 14, 1094–1104
8. Liu, Y., and Eisenberg, D. (2002) Protein Sci. 11, 1285–1299
9. Ross, C. A. (2002) Neuron 35, 819–822
10. Zoghbi, H. Y., and Orr, H. T. (2000) Annu. Rev. Neurosci. 23, 217–247
11. Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenberg, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) Cell 90, 549–558
12. Huang, C. C., Faber, P. W., Persichetti, F. Mittal, V., Vonsattel, J. P., McDonald, M. E., and Gussella, J. F. (1998) Somat. Cell Mol. Genet. 24, 217–233
13. Chen, S., Berthelier, V., Yang, W., and Wetzel, R. (2001) J. Mol. Biol. 311, 173–182
14. Bucicchantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zard, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002) Nature 416, 507–511
15. Nucifora, F. C., Jr., Sasaki, M., Peters, M. F., Huang, H., Cooper, J. K., Yamada, M., Takahashi, H., Tsui, S., Troncoso, J., Dawson, V. L., Dawson, T. M., and Ross, C. A. (2001) Science 291, 2423–2428
16. Scherzinger, E., Breuer, P., Boteva, R., Behrends, C., Tsvetkov, N., Strippe, N., Sakahira, H., Siegers, K., Hayer-Hartl, M., and Hartl, F. U. (2000) Mol. Cell 15, 95–105
17. Nilsberth, C., Westland-Danielsson, A., Eckman, C. B., Condon, M. M., Axelman, K., Forsell, C., Sten, C., Luthman, J., Teplow, D. B., Younkim, S. G., Naslund, J., and Lannfelt, L. (2001) Nat. Neurosci. 4, 887–893
18. Bates, G. (2003) Lancet 361, 1642–1644
19. Poitier, M. A., Li, H., Macosko, J., Cai, S., Amsel, M., and Ross, C. A. (2002) J. Biol. Chem. 277, 41032–41037
20. Perutz, M. F., Pope, B. J., Owen, D., Wanker, E. E., and Scherzinger, E. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 5596–5600
21. Tanaka, M., Machida, Y., Nishikawa, Y., Akagi, T., Hashikawa, T., Fujitawa, J., and Nukina, N. (2003) J. Biol. Chem. 278, 34717–34724
22. Chen, S., Berthelier, V., Hamilton, J. B., O’Nuallain, B., and Wetzel, R. (2002) Biochemistry 41, 7391–7399
23. Scherzinger, E., Sitaller, A., Schweiger, K., Heiser, V., Lurz, R., Hasenberg, R., Bates, G. P., Lehrach, H., and Wanker, E. E. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 4604–4609
24. Harper, J. D., and Lansbury, P. T., Jr. (1997) Annu. Rev. Biochem. 66, 385–407
25. Dobson, C. M. (2003) Nature 426, 884–890
26. Dobson, C. M. (2004) Semin. Cell Dev. Biol. 15, 3–16
27. Schymkowitz, J. W., Rousseau, F., and Itzhaki, L. S. (2000) J. Mol. Biol. 301, 199–204
28. Seeliger, M. A., Schymkowitz, J. W., Rousseau, F., Wilkinson, H. R., and Itzhaki, L. S. (2002) Biochemistry 41, 1202–1210
29. Ilag, L. L., Westblade, L. F., Deshayes, C., Kolb, A., Busby, S. J., and Robinson, C. V. (2004) Structure 12, 269–275
30. Spackman, D. H., Stein, W. H., and Moore, S. (1958) Anal. Chem. 30, 1190–1206
31. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) J. Mol. Biol. 273, 729–739
32. Gurrera, J. L., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 4224–4228
33. Fandrich, M., and Dobson, C. M. (2002) EMBO J. 21, 5682–5690
34. Perutz, M. F., Finch, J. T., Berriman, J., and Lesk, A. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 5591–5595
Fibril Formation in Cks1

35. Sikorski, P., and Atkins, E. (2005) Biomacromolecules 6, 425–432
36. Venyaminov, S., and Kalnin, N. N. (1990) Biopolymers 30, 1243–1257
37. Rahmelow, K., Hubner, W., and Ackermann, T. (1998) Anal. Biochem. 257, 1–11
38. Krull, L. H., Wall, J. S., Zobel, H., and Dimler, R. J. (1965) Biochemistry 4, 626–633
39. Bevivino, A. E., and Loll, P. J. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 11955–11960
40. Arvai, A. S., Bourne, Y., Ficici, M. J., and Tainer, J. A. (1995) J. Mol. Biol. 249, 835–842
41. Khazanovich, N., Bateman, K., Chernia, M., Michalak, M., and James, M. (1996) Structure 4, 299–309
42. Chen, Y. W., Stott, K., and Perutz, M. F. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 1257–1261
43. Sambashivan, S., Liu, Y., Sawaya, M. R., Gingery, M., and Eisenberg, D. (2005) Nature 437, 266–269
44. Perutz, M. F., and Windle, A. H. (2001) Nature 412, 143–144
45. Jenkins, J., and Pickersgill, R. (2001) Prog. Biophys. Mol. Biol. 77, 111–175
46. Wetzel, R. (2002) Structure 10, 1031–1036
47. Kishimoto, A., Hasegawa, K., Suzuki, H., Taguchi, H., Namba, K., and Yoshida, M. (2004) Biochem. Biophys. Res. Commun. 315, 739–745
48. Bouchard, M., Zardo, J., Nettleton, E. J., Dobson, C. M., and Robinson, C. V. (2000) Protein Sci. 9, 1960–1967
49. Khurana, R., and Fink, A. L. (2000) Biophys. J. 78, 994–1000
50. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R., and Eisenberg, D. (2005) Nature 435, 773–778
51. Wellington, C. L., and Hayden, M. R. (2000) Clin. Genet. 57, 1–10
52. Sugars, K. L., and Rubinsztein, D. C. (2003) Trends Genet. 19, 233–238
53. Perez, M. K., Paulson, H. L., Pendse, S. J., Saionz, S. J., Bonini, N. M., and Pittman, R. N. (1998) J. Cell Biol. 143, 1457–1470
54. Kazantiev, A., Preisinger, E., Dranovsky, A., Goldgaber, D., and Housman, D. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 11404–11409
55. McCampbell, A., Taylor, J. P., Tate, A. A., Rohitschek, J., Li, M., Walcott, J., Merry, D., Chai, Y., Paulson, H., Sobae, G., and Fischbeck, K. H. (2000) Hum. Mol. Genet. 9, 2197–2202
56. Bence, N. F., Sampat, R. M., and Kopito, R. R. (2001) Science 292, 1552–1555
57. Venkatraman, P., Wetzel, R., Tanaka, M., Nukina, N., and Goldberg, A. L. (2004) Mol. Cell 14, 95–104
58. Stott, K., Blackburn, J. M., Butler, P. J., and Perutz, M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 6509–6513
59. Tanaka, M., Machida, Y., Nishikawa, Y., Akagi, T., Morishima, I., Hashikawa, T., Fujisawa, T., and Nukina, N. (2002) Biochemistry 41, 10277–10286
60. Masino, L., Kelly, G., Leonard, K., Trostier, Y., and Pastore, A. (2002) FEBS Lett. 513, 267–272