The Residual Pro-Part of Cathepsin C Fulfills the Criteria Required for an Intramolecular Chaperone in Folding and Stabilizing the Human Proenzyme[†]

Blaž Cigić,*[,‡,§] Søren W. Dahl,[||] and Roger H. Pain[‡]

Department of Biochemistry and Molecular Biology, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia, and Unizyme Laboratories, Dr Neergaads vej 17, DK-2970 Horsholm, Denmark

Received April 18, 2000; Revised Manuscript Received July 26, 2000

ABSTRACT: The 13.5 kDa N-terminal part of the propeptide remains associated with mature cathepsin C after proteolytic activation and excision of the activation peptide. This residual pro-part, isolated from the recombinant enzyme, folds spontaneously and rapidly to a stable, compact monomer with secondary structure and stable tertiary interactions. Folding and unfolding kinetics of the residual pro-part with intact disulfides are complex, and accumulation of transient intermediates is observed. The cleaved form of the pro-part isolated from natural human cathepsin C also folds, suggesting that the intact form comprises two folding domains. The linkages of the two disulfide bridges have been established as 30−118 and 54−136 for the native enzyme. The native disulfide bonds can be re-formed from the fully reduced and denatured state by oxidative refolding, resulting in a domain that is spectroscopically indistinguishable from the original refolded residual pro-part. Both disulfides are solvent-exposed and can be reduced in the absence of denaturant. The reduced form retains most or all of the native tertiary structure and is only $\approx$2 kcal·mol^{-1} less stable than the oxidized form. It folds fast relative to the rate of biosynthesis, to the same conformation as the oxidized form. Folding and disulfide formation are sequential. These results indicate that the proenzyme folds sequentially in vivo and that the residual pro-part constitutes a rapidly and independently folding domain that stabilizes the mature enzyme. It thus fulfills the criteria required of an intramolecular chaperone. It may also be involved in stabilizing the tetrameric structure of the mature enzyme.

Proteinases are synthesized as zymogens that are activated by proteolytic cleavage with the loss of from 2 or 3 to more than 200 residues. Their propeptides are implicated in the regulation of proteolysis and in sorting zymogens to their final location in the cell where the activation process takes place. The majority of propeptides are larger than 50 residues, and these are frequently considered as assisting in the folding of the catalytic part.

The involvement of propeptides in folding has been most extensively studied with the bacterial serine proteinases α-lytic protease and subtilisin. Coexpression of the propeptide with the mature enzyme was shown to be essential to enable folding of the latter in vivo (*1*, *2*). In vitro, α-lytic protease (*3*) and subtilisin (*4*) fold to a molten globule-like intermediate state, further folding to the native state requiring addition of the propeptide. The mature enzyme is thus in a metastable conformation where the completely unfolded and intermediate states are more stable than the proteolytically active enzyme (*5*). α-Lytic protease is, however, kinetically stable because of the large activation barrier between the folded and unfolded states. This barrier is lowered in the folding process by the propeptide which stabilizes the transition state (*5*). Subsequent proteolytic degradation then locks the α-lytic protease in its metastable active state. Recently, a model has been suggested in which α-lytic protease folds in vitro in a sequential manner, the propeptide folding first, then being able to catalyze folding of the protease domain (*6*).

Sequential folding is even more feasible in vivo, since individual domains can fold co-translationally. Early experiments on the biosynthesis of antibody fragments in cell-free systems (*7*) showed that an immunoglobulin domain, complete with its disulfide bond, can fold co-translationally and independently of the rest of the protein. Sequential folding during in vitro translation has also been observed for firefly luciferase (*8*, *9*). The in vitro folding of tailspike endorhamnosidase from salmonella bacteriophage P22 proceeds through rate-limiting intermediates with the same folding rates as in vivo folding, suggesting sequential folding of domains (*10*, *11*).

Propeptides of cysteine zymogens bind across the active site cleft, inhibiting the enzyme (*12*). Mutation experiments on the complete propeptide have shown that it is necessary for folding and/or processing of the proenzyme (*13*−*15*). Moreover, independent evidence has shown that the proregion of *Bombyx mori* cysteine proteinase is, like that of α-lytic protease, essential for enabling the catalytic domain to fold to the active conformation (*16*).

[†] This work was supported by the Ministry of Science and Technology of the Republic of Slovenia.

* To whom correspondence should be addressed. Phone: +386 1 423 11 61. Fax: +386 1 256 62 96. E-mail: blaz.cigic@bf.uni-lj.si.

[‡] Jožef Stefan Institute.

[§] Present address: Department of Food Science and Technology, Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, 1000 Ljubljana, Slovenia.

[||] Unizyme Laboratories.

10.1021/bi0008837 CCC: $19.00
Published on Web 09/14/2000

Propeptides of the papain-like enzymes are usually not larger than 100 amino acids and are completely cleaved out during proteolytic activation. The known exceptions are cathepsin F and cathepsin C, with propeptides of 251 and 206 residues, respectively (*17*, *18*). The sequence of cathepsin F propeptide indicates the presence of 3 distinct modular domains: an N-terminal sequence predicted to adopt a cystatin-like fold, a 50 residue linking peptide, and a C-terminal sequence with similarity to the cathepsin L propeptide and therefore presumably responsible for suppressing catalytic activity in the proenzyme. In the precursor of cathepsin C (dipeptidyl aminopeptidase I, EC 3.4.14.1), the C-terminal 10 kDa sequence of the propeptide shows similarities to the propeptide of cathepsin L (*19*) and is excised on activation, while the N-terminal 13.5 kDa part, referred to here as the residual pro-part, is unusual in remaining associated by strong noncovalent interactions (*20*) as part of the mature enzyme (*21*, *22*). Its sequence shows no similarity to other proteinases or their precursors in the SwissProt and Protein Data Banks, so it is therefore of interest to question whether it plays a specific role in the zymogen and the mature enzyme. The other unusual feature of cathepsin C is that the proenzyme rapidly oligomerizes in vivo, followed by activation to a tetrameric mature enzyme (*21*, *23*).

A related question concerns whether, in vivo, the procathepsin C sequence codes for independent folding domains which, like α-lytic protease, fold sequentially, or whether folding of the N-terminal region depends on that of the C-terminal region, as in the case of *β*-lactamases (*24*).

We therefore isolated the bound, residual pro-part from cathepsin C, characterized it chemically and spectroscopically, and investigated the reversibility, equilibria, and kinetics of its folding and unfolding. It is shown to fold spontaneously and independently of the remainder of the propeptide, and to be capable of folding before the catalytic moiety of the proenzyme is synthesized in vivo. It thus satisfies the criteria necessary for assisting the folding of the proenzyme.

EXPERIMENTAL PROCEDURES

Materials. Enzymically active, recombinant human dipeptidyl aminopeptidase I (DPPI, cathepsin C) was obtained by heterologous expression of the proenzyme in insect cells and, following activation, was purified and characterized (S. W. Dahl, personal communication). TPCK-treated trypsin for sequential analysis was from Sigma. GdmCl,[1] 5,5′-dithiobis-(2-nitrobenzoic acid), and 8-anilino-1-naphthalenesulfonic acid were from Serva and KI from Aldrich.

Isolation of the Residual Pro-Part. Recombinant cathepsin C was dissolved in 200 mM phosphate, pH 7, containing 6 M GdmCl and stood for 1 h at room temperature. Refolding was started with 1:5 dilution of denatured protein into 200 mM phosphate, pH 7. The cathepsin C concentration during refolding was 0.25 mg/mL or lower. After 30 min, insoluble aggregates were removed by centrifugation at 10000*g* for 10 min. The soluble proteins were concentrated by ultrafiltration (Amicon YM-10) to 10% of the volume followed by further centrifugation. The residual pro-part was separated from the soluble aggregates on a calibrated Superdex S-200 column equilibrated with 100 mM phosphate/300 mM NaCl, pH 7. The concentration of the residual pro-part was determined from the absorbance at 280 nm using $A^{1mg/mL}_{280nm} = 2.39$ (*20*).

Characterization of the Residual Pro-Part. The fluorescence spectra of the residual pro-part were measured with excitation at 280 nm using a PTI Fluorescence System version 1.2X. The spectra were scanned in 200 mM phosphate buffer, pH 7, in the presence and absence of 6 M GdmCl, 20 mM DTT, at a protein concentration of 5 *μ*g/mL. CD spectra were recorded on an AVIV CD Spectrometer 62A DS in 100 mM phosphate buffer, pH 7, in the presence and absence of 20 mM DTT. ANS fluorescence spectra were measured with excitation at 370 nm, with 50 *μ*mol/L ANS in the presence of a 2 *μ*mol/L sample of the residual pro-part dissolved in 100 mM phosphate, pH 7. SDS−PAGE was performed on a Phast System apparatus (Pharmacia, Sweden) using 8−25% gradient gels following the manufacturer's instructions. N-Terminal amino acid sequence analyses employed a Procise 492 Applied Biosystems amino acid sequenator system.

Equilibrium Studies of GdmCl-Induced Unfolding and Refolding. Changes in protein structure were detected by recording the tryptophan fluorescence at 320 nm with excitation at 280 nm and ellipticity in the far-UV region at 220 nm. Experiments were performed by titration with 20 min equilibration at each concentration of the denaturant in 100 mM phosphate, pH 7, in the presence and absence of 20 mM DTT. This equilibration time is adequate, since the half-time for the slow phase of unfolding in 4 M GdmCl is 100 s and of refolding in 3.2 M GdmCl is 5 s. In 3 M GdmCl, 90% of the slow phase of unfolding (itself only 10% of the total change) would take place in 500 s, meaning that any difference from the equilibrium value would be well within experimental error. The observed coincidence of the folding and unfolding curves supports this conclusion. All experiments were performed at 25 °C. Starting concentrations of the residual pro-part for the fluorescence and CD measurements, both in a 10 mm cell, were 10 and 20 *μ*g/mL, respectively. The results were corrected for dilution and analyzed according to the equation for a two-state transition with baseline slope correction (*25*):

$$S_{obs} = \{S_n + n_n[D] + (S_u + n_u[D])\exp[(m[D] - \Delta G^\circ_{N-U})/RT]\}/\{1 + \exp[(m[D] - \Delta G^\circ_{N-U})/RT]\} \quad (1)$$

S_{obs} is the measured signal, S_n and S_u are the values for the native and the unfolded protein at zero denaturant, respectively, [D] is the denaturant concentration, and n_n and n_u are the slopes of pre- and posttranslational baselines, respectively. ΔG°_{N-U} is the difference in Gibbs free energy between native and unfolded protein in the absence of denaturant, and $m = \delta\Delta G_{N-U}/\delta[D]$.

Unfolding and Refolding Kinetics of the Oxidized Residual Pro-Part. The kinetics of change of the intrinsic tryptophan fluorescence were monitored using a DX.17MV sequential stopped-flow spectrofluorometer with excitation at 280 nm and the emission cutoff filter at 335 nm. Unfolding was initiated by rapid mixing of 1 volume of the protein solution with 10 volumes of 100 mM phosphate, pH 7, containing

[1] Abbreviations: DTT, dithiothreitol; GdmCl, guanidinium chloride; GSH, reduced glutathione; GSSG, oxidized glutathione.

various concentrations of denaturant. For refolding, the protein was first denatured by 30 min incubation in 6 M GdmCl and then mixed with 10 parts of the refolding buffer.

The kinetics of change of quenching of tryptophan fluorescence were measured in the presence of 150 mM KI and 0.1 mM $Na_2S_2O_3$. Folding and unfolding were initiated by rapid 1:5 mixing of protein with 100 mM phosphate, pH 7, containing GdmCl, following the emission above 295 nm with excitation at 280 nm. Fluorescence measurements in the presence and absence of the quencher were performed at a protein concentration of 5 μg/mL. Extrapolation of the signal for the folded residual pro-part to the unfolding conditions allowed estimation of the amplitudes for each phase observed in the unfolding. The signal corresponding to the unfolded residual pro-part under folding conditions, obtained by extrapolation from the denatured state, allowed the amplitudes in the refolding kinetics to be estimated.

The kinetics of secondary structure formation and unfolding were determined from the change of ellipticity at 220 nm. Reactions were initiated by manual mixing of protein and buffer, with a dead time of 7 s. Refolding in the presence of 2.5 M GdmCl and unfolding in 5 M GdmCl were performed at a residual pro-part concentration of 60 μg/mL in a 1 cm cell at 10 °C.

All refolding reactions were analyzed according to a single-exponential equation and unfolding reactions as the sum of two single-exponential reactions.

Refolding of the Reduced Residual Pro-Part. The residual pro-part was dissolved in 6 M GdmCl, 0.5 M Tris, pH 8, containing 50 mM DTT. After an hour, the sample was applied to a size exclusion chromatography column equilibrated with 6 M GdmCl, 150 mM AcOH, pH 2.5, to remove DTT. Folding was started by 1:10 dilution of the reduced residual pro-part into the refolding buffer to a final concentration of 10 μg/mL. The refolding buffer contained 3 M urea, 0.6 M GdmCl, 2.5 mM GSH, 0.25 mM GSSG, and 0.3 M Tris, pH 7. The products of the refolding were analyzed by a size exclusion chromatography column equilibrated with 300 mM NaCl and 100 mM phosphate, pH 7.

Colorimetric Quantitation of Free Sulfhydryls. Free sulfhydryl content was quantitatively determined with Ellman's reagent, 5,5′-dithiobis(2-nitrobenzoic acid), measuring the absorbance at 412 nm of the formed thionitrobenzoate (*26*). The residual pro-part at a concentration 5 μmol/L was dissolved in 200 mM Tris, pH 8, containing 5 M GdmCl and 200 μmol/L of Ellman's reagent in a total volume of 0.4 mL. The absorbance was measured after 10 min incubation. The number of free cysteines in the residual pro-part was determined from the calibration curve obtained with free cysteine in a concentration range 3−30 μmol/L.

The Position of Disulfide Bridges. The location of disulfide bridges was determined by classical fragmentation strategy (*27*). Residual pro-part (20 μmol/L) in 200 mM Tris, pH 7.5, was digested with 2 μmol/L trypsin for 2 h at 37 °C. The reaction was stopped with boiling at 100 °C for 5 min, half of the sample in the presence of 50 mM DTT and the other half in the absence of reducing agent. The resulting peptide mixtures were separated by HPLC (Milton Roy) on a Chrompack C18 column (300 × 3 mm) with a gradient of acetonitrile in an aqueous buffer containing 0.1% TFA. The peptides containing a disulfide bond were identified by comparison of the reduced and oxidized residual pro-parts, collected, reduced, and re-run on the column. N-Terminal amino acid sequence analyses of the disulfide-bonded peptides were performed using a Procise 492 Applied Biosystems amino acid sequenator system.

RESULTS

The Residual Pro-Part of Cathepsin C. Unlike the major proportion of the residual pro-part present in cathepsin C in human tissue (*20*−*23*), the peptide isolated from the recombinant enzyme is uncleaved and homogeneous, as shown by reducing SDS−PAGE and confirmed by N-terminal amino acid analysis (data not shown).

The residual pro-part in mature cathepsin C is bound to the rest of the enzyme solely by noncovalent interactions (*20*). The binding is strong since, in 2 M GdmCl, only the tetramer with an apparent molecular mass of 180 kDa is observed when analyzed by size exclusion chromatography. The higher concentrations of GdmCl required to separate the residual pro-part from the enzyme lead to simultaneous denaturation of all the constituent peptides, and to the irreversible loss of oligomeric structure and activity (*28*). Dilution of the denatured cathepsin C into aqueous buffer leads to substantial aggregation. Attempts to optimize refolding by staggered dilution or by the presence of nonionic or zwitterionic detergents did not result in any recovery of activity. This failure to refold in vitro may be due to the cleavage, during activation, of the catalytic moiety and/or to the absence of the excised propeptide (*20*−*23*).

Size exclusion chromatography of the products of the attempted refolding showed the presence of a soluble fraction with a symmetrical peak. The single N-terminal amino acid sequence found, DTPAX, is that of the residual pro-part. Reducing SDS−PAGE gave a single band corresponding to an apparent molecular mass of ∼25 kDa, the same as for the uncleaved residual pro-part isolated from human tissue (*20*). This shows that the residual pro-part in the recombinant enzyme is, like the human enzyme (*20*), highly glycosylated, the mass of the protein chain being 13.5 kDa. It was shown by mutational analysis to be glycosylated at all three sites, with little difference in the nature of the glycosylation from that of the natural enzyme (S. W. Dahl, personal communication).

The positions of the disulfide bonds in the residual pro-part isolated from the active recombinant enzyme were determined. Disulfide-containing peptides resulting from tryptic cleavage were separated on reverse phase HPLC. On reduction, one gave peptides with sequences DVXCSVMGPQE and NWACFTG, indicating that cysteines 54 and 136 are disulfide bonded. The second pair gave N-terminal sequences DTPA and VTTY, confirming that the other disulfide bond is formed between cysteines 30 and 118.

Characterization of the Folded State of the Residual Pro-Part of Cathepsin C. The fluorescence of the soluble residual pro-part has a maximum at 325 nm, compared with the peak at 350 nm observed for the GdmCl denatured state (Figure 1). A substantial proportion of the aromatic residues are therefore shielded from solvent.

Near-UV circular dichroism provides a sensitive fingerprint of tertiary interactions and of the dynamics of a protein. The residual pro-part has a well-defined near-UV CD

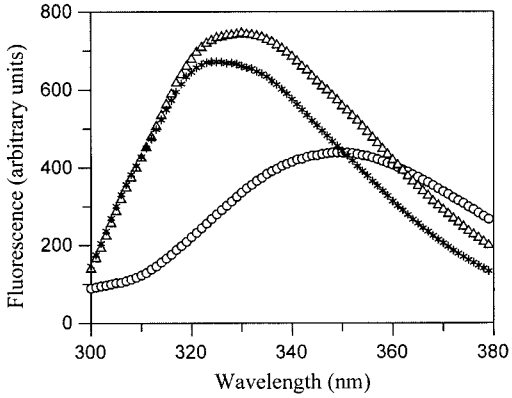


FIGURE 1: Fluorescence emission spectra of the residual pro-part isolated from mature, recombinant cathepsin C. Spectra were scanned in the presence of 6 M GdmCl (○) and 20 mM DTT (△) and in the absence of reducing agent and denaturant (*).

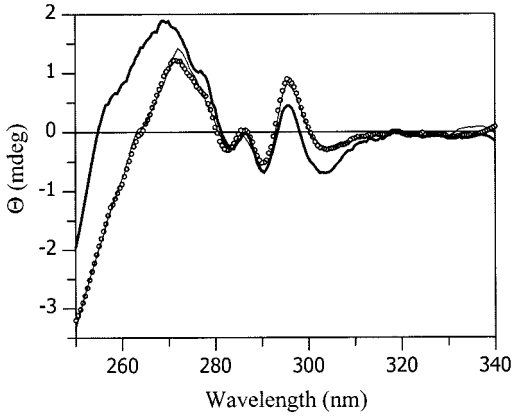


FIGURE 2: Near-UV CD spectra of the recombinant residual pro-part in the absence (thin line) and in the presence (thick line) of 20 mM DTT. (○) The reoxidized residual pro-part refolded from the reduced and denatured state.

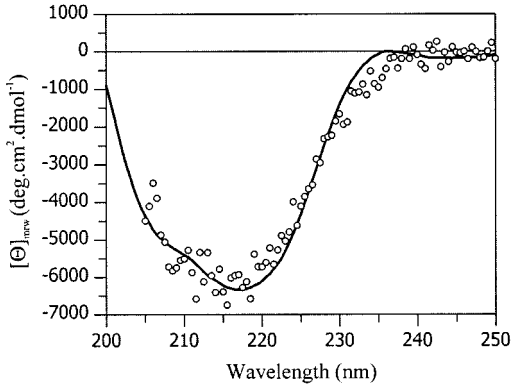
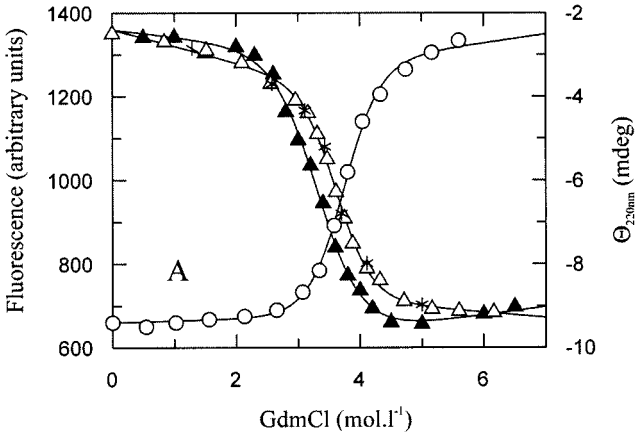


FIGURE 3: Far-UV CD spectra of the recombinant residual pro-part in the absence (line) and in the presence (○) of 20 mM DTT.



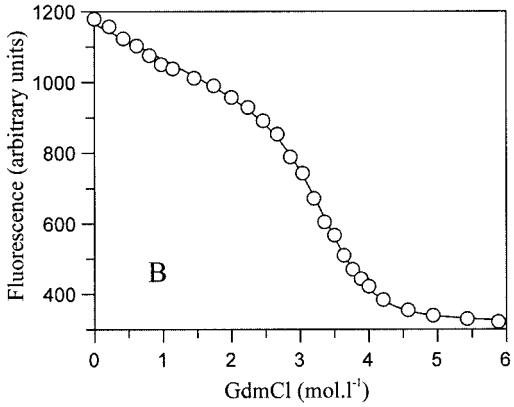


FIGURE 4: (A) GdmCl-induced unfolding transition of the oxidized residual pro-part from recombinant cathepsin C monitored by the fluorescence change at 320 nm (△) and by the decrease of the far-UV ellipticity at 220 nm (○); (*) refolding from 5 M GdmCl measured by the fluorescence change. Unfolding of the pro-part from natural human cathepsin C monitored by the fluorescence change at 320 nm (▲). (B) GdmCl-induced unfolding transition of the reduced residual pro-part monitored by the fluorescence change at 320nm (○). The lines in all cases represent analyses of the data according to a two-state model (eq 1).

Table 1: Parameters Characterizing the Equilibrium Unfolding of the Residual Pro-Part of Cathepsin C

	ΔG°_{N-U} (kcal·mol^{-1})	m (kcal·mol^{-1}·M^{-1})	C_m[a] (M)
[1]ox. CD	7.6 ± 0.5	2.0 ± 0.1	3.7 ± 0.1
[1]ox. fluor.	7.6 ± 0.4	2.1 ± 0.1	3.6 ± 0.1
[1]red. fluor.	5.7 ± 0.2	1.7 ± 0.1	3.4 ± 0.1
[2]ox. fluor.	5.6 ± 0.3[b]	1.7 ± 0.1[b]	3.3 ± 0.2[b]

[a] C_m is the concentration of GdmCl at the midpoint of the transition: $C_m = \Delta G^{\circ}_{N-U}/m$. Abbreviations: ox., oxidized; red., reduced; fluor., fluorescence probe. Origin of the residual pro-part was [1]recombinant and [2]natural human cathepsin C. [b]Apparent values, due to the heterogeneity of the peptide (see text).

spectrum (Figure 2), with contributions from tryptophan, tyrosine, and phenylalanine, confirming that persistent tertiary interactions are established. In addition, the absence of any enhancement of ANS fluorescence shows that the latter is not bound by the soluble residual pro-part, confirming that it is not in a partially folded, molten globule-like state. In addition, the far-UV CD spectrum shows well-developed secondary structure, with major contributions of β-sheet and aperiodic conformation (Figure 3).

Additional evidence that the residual pro-part in aqueous buffer is folded is provided by equilibrium denaturation experiments (Figure 4A). It unfolds in a thermodynamically reversible manner, and the coincidence of the curves monitored by fluorescence and far-UV CD is characteristic of a two-state transition. Further, the data fit well to a two-state model (Figure 4A), resulting in a thermodynamic stability, $\Delta G^{\circ} = 7.7$ kcal·mol^{-1}, characteristic of a stable globular protein, and an m value of 2.0 kcal·mol^{-1}·M^{-1} (Table 1), indicating typical cooperativity.

Kinetics of Renaturation of the Oxidized Residual Pro-Part. Refolding was followed on the stopped-flow time scale by the decrease of fluorescence intensity above 335 nm.

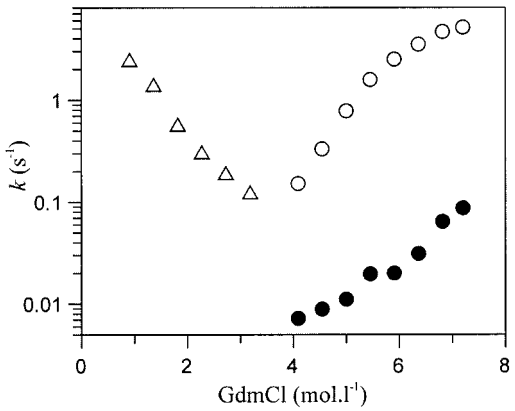


FIGURE 5: Apparent rate constants, as a function of GdmCl concentration, for the unfolding (○,●) and refolding (△) of the recombinant residual pro-part obtained by stopped-flow fluorescence measurements.

Table 2: Apparent Rate Constants and Amplitudes of Folding of the Residual Pro-Part in 2.5 M GdmCl and Unfolding in 5 M GdmCl

	refolding		unfolding			
	slow phase		fast phase		slow phase	
	rate (s^{-1})	amplitude (%)	rate (s^{-1})	amplitude (%)	rate (s^{-1})	amplitude (%)
far-UV CD[a]	0.048	30	0.052	40	0.0012	40
fluorescence[b]	0.23	15	0.8	25	0.012	25
I^- quenching[b]	0.18	70	1.1	50	0.017	50

[a] Experiments were performed at 10 °C. [b] Experiments were performed at 25 °C.

Approximately 85% of the change of fluorescence occurs in the burst phase of folding. The subsequent reaction is monophasic, and the log rate constants are inversely proportional to the denaturant concentration over the concentration range 0.5−3.2 M GdmCl (Figure 5). The overall folding of the residual pro-part is very fast, extrapolation of the rate constants to zero GdmCl concentration leading to a half-time of 0.1 s at 25 °C.

Formation of a compact state during folding was monitored by following the dequenching of aromatic fluorescence by iodide ion. One kinetically resolved phase with a folding rate constant of 0.18 s^{-1} in the presence of 2.5 M GdmCl represented 70% of the total amplitude, as established from the extrapolation of the fluorescence of the unfolded state to the folding conditions. The first 30% of dequenching occurred in the burst phase, showing that most of the process of sequestering aromatic residues from solvent occurs in the later stages of folding (Table 2).

Similar kinetic phases were observed for the formation of secondary structure. The change of ellipticity at 220 nm was measured at 10 °C to allow folding in 2.5 M GdmCl to be observed following manual mixing. As for the intrinsic fluorescence, the greater proportion of native ellipticity was regained in the burst phase, followed by a slower phase with a rate constant consistent with that for fluorescence under the same solvent conditions (Figure 7B and Table 2). The progression of ANS binding is complementary to these results. Stopped-flow fluorescence showed that ANS is bound early on in folding, although much less than to other proteins (*29*). The desorption of ANS at 10 °C takes place over the same time range as the slow phases of changes in tryptophan

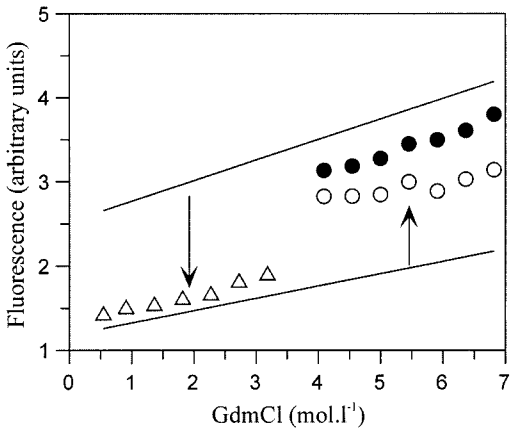


FIGURE 6: Amplitudes of the burst phases and of the observed phases in the unfolding (○,●) and refolding (△) of the recombinant residual pro-part. The upper line refers to the signal of the unfolded and the bottom line to the signal of the folded residual pro-part. The arrows show the direction of change.

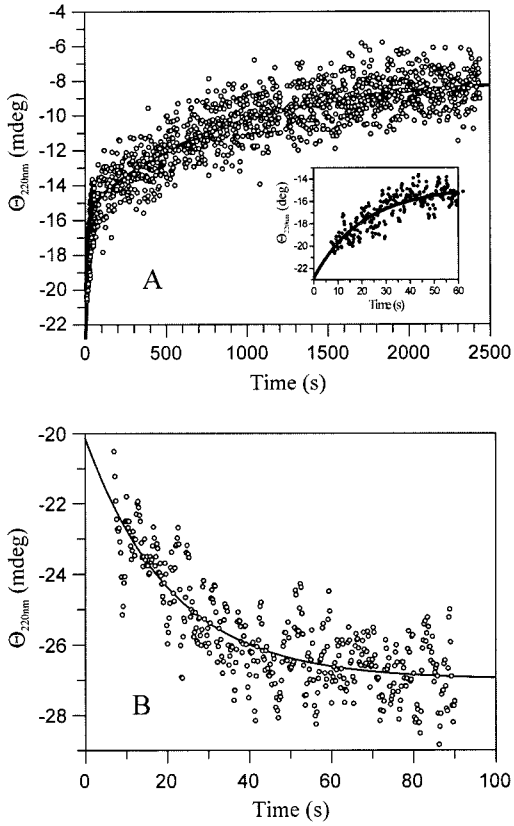


FIGURE 7: Rates of unfolding (A) and refolding (B) of the recombinant residual pro-part measured by the change of ellipticity at 220 nm. The solid lines are the best fits of the experimental data to (A) the sum of two single-exponential functions and (B) a single-exponential function.

fluorescence and far-UV ellipticity, indicating the accumulation of an intermediate with some molten globule-like characteristics at a relatively late stage in folding.

Kinetics of Denaturation of the Oxidized Residual Pro-Part. The kinetics of unfolding were studied using intrinsic fluorescence, fluorescence quenching, and far-UV CD. The rate of increase of fluorescence shows that unfolding is not a simple two-state transition from the native to the unfolded state (Figure 5). Two kinetically resolved phases are observed in the range from 4 to 7 M GdmCl, the rate constants in 5

Table 3: Comparison of Unfolding and Refolding Kinetics for Residual Pro-Part from 'Natural' and Recombinant Cathepsin C[a]

	unfolding		refolding	
	d log k_u/ d[GdmCl]	log k_u° (s^{-1})	d log k_f/ d[GdmCl]	log k_f° (s^{-1})
rec. phase 1	0.51	−2.7	−0.6	0.9
rec. phase 2	0.36	−3.7	−	−
nat. phase 1	0.53	−2.7	−0.8[b]	1.5[c]
nat. phase 2	0.58	−4.4	−	−

[a] Kinetics measured by fluorescence; see Experimental Procedures. k° is the rate constant extrapolated to zero concentration of GdmCl. [b] The slope is linear between 2 and 3.3 M GdmCl, but nonlinear at lower concentrations, extrapolating to a value for log k_f° of between 0.1 and 0.4 s^{-1}. [c] Extrapolation of the linear part of the curve to zero denaturant concentration.

M GdmCl being 0.8 and 0.012 s^{-1}. However these two phases constitute only about 50% of the total amplitude (Figure 6), the remainder of the change in fluorescence occurring within the burst phase. The dependence of the amplitude of the latter on denaturant concentration (Figure 6) suggests that this 'phase' represents the effect of change of solvent.

When the decrease of fluorescence was monitored in the presence of iodide, two phases accounting for the total amplitude were observed (Table 2). These quenching experiments were performed under conditions (cutoff filter 295 nm) such that fluorescence in the absence of quencher was unchanged during unfolding. The observed change in the presence of quencher thus reflects only the solvent exposure of aromatic residues during unfolding.

Binding of ANS was not observed in either phase, indicating that any accumulating intermediates are not molten globule-like.

Loss of secondary structure in 5 M GdmCl was monitored by change of ellipticity at 220 nm at 10 °C, following manual mixing (Figure 7A). Again, two kinetically resolved phases were observed, with rate constants of 0.052 and 0.0012 s^{-1}, accounting for 80% of the total amplitude (Table 2). The other 20% was lost within the dead time of mixing (7 s).

Comparison of the Residual Pro-Parts from Natural and Recombinant Human Cathepsins C. To confirm that results on the folding behavior of the residual pro-part from recombinant enzyme can be related to the 'natural' proenzyme, certain experiments were carried out on the residual pro-part from the natural human enzyme, bearing in mind that approximately 50% of this peptide is cleaved at residue 58 or 61 (*20*), although still joined by disulfide bonds. The equilibrium unfolding transition (Figure 4A) was similar to that of the homogeneous recombinant peptide, the apparent values of free energy, *m*, and C_m being slightly lower than those of the recombinant peptide (Table 1). This can be attributed to the heterogeneity of the natural pro-part, the cleaved fraction being apparently slightly less stable than the intact fraction.

The kinetics of folding and unfolding, as a function of GdmCl concentration, follow the same pattern for both sources of pro-part (Figure 5). The unfolding kinetics exhibit two phases, each of which has rate constants of the same order for the two pro-parts (Table 3). The folding kinetics for the natural pro-part show a nonlinear dependence of rate constant on denaturant concentration, indicative of heterogeneity. At higher concentrations, where the more stable intact form from the recombinant enzyme will be the main folding species, the plot extrapolates to a value close to that for the recombinant pro-part (Table 3).

Properties of the Reduced Residual Pro-Part in the Absence of Denaturant. To discover whether the two disulfide bonds are important for the stability and folding of the residual pro-part under physiological conditions, the reduced protein was characterized. After incubation for an hour in 200 mM phosphate, pH 7, containing 20 mM DTT, the residual pro-part was acidified to pH 3 and the reducing agent removed on a size exclusion chromatography column equilibrated with 6 M GdmCl at pH 2.5. The number of free cysteines, produced by reduction in the absence of denaturing agent, was determined using Ellman's reagent to be 3.95 residues per molecule. The oxidized residual pro-part and the form reduced in the presence of 6 M GdmCl possessed 0.05 and 3.95 free cysteines, respectively. Both disulfide bonds can thus be reduced in the absence of the denaturant.

The far-UV CD spectrum shows that the secondary structure is not significantly different from that in the oxidized form (Figure 3), the small difference in ellipticity around 235 nm being attributable to the dithiol chromophore. More significantly, most of the tertiary interactions and dynamics involving aromatic residues are conserved in the reduced state, as shown by the near-UV CD spectrum (Figure 2). The fluorescence spectrum of the reduced residual pro-part in the presence of 20 mM DTT differs from that of the oxidized residual pro-part (Figure 1). The increase of λ_{max} from 325 to 330 nm and of fluorescence intensity at wavelengths above 320 nm indicates that the environment of one or more of the tryptophans has changed on reduction. This, however, could be due to loss of quenching by a neighboring disulfide bond. Reduction thus has little, if any, effect on the conformation of the residual pro-part.

The thermodynamic stability of the reduced residual pro-part was determined from equilibrium unfolding experiments monitored by fluorescence at 320 nm (Figure 4B). Like the oxidized form, the equilibrium unfolding is fitted well by a two-state model, the derived thermodynamic parameters, $\Delta G^\circ_{N-U} = 5.7$ kcal·mol^{-1} and $m = 1.7$ kcal·mol^{-1}·M^{-1} (Table 1), showing that the reduced residual pro-part is less stable than the fully oxidized form by just 2 kcal·mol^{-1} and that the transition is slightly less cooperative.

Oxidative Refolding of the Reduced and Denatured Residual Pro-Part. Initial attempts to refold and reoxidize the reduced and denatured residual pro-part resulted in strong aggregation. It was necessary to find conditions that would reduce the likelihood of intermolecular disulfides forming, while allowing the protein to fold and make the correct intramolecular bonds. A refolding solvent of 3.15 M urea and 0.6 M GdmCl was found to be optimal.

Residual pro-part was fully reduced with DTT in the presence of GdmCl and applied to a size exclusion chromatography column equilibrated with 6 M GdmCl at pH 2.5 in order to remove the reducing agent and stabilize the sulfhydryls in the fully unfolded pro-part. Renaturation was started with a 1:10 dilution of this denatured protein into a buffer containing 3.5 M urea and the redox pair 2.5 mM GSH and 0.25 mM GSSG. After 8 h, the samples were analyzed by size exclusion chromatography. More than 95%

of the protein was eluted at the same volume as for the unreduced peptide isolated from the mature enzyme.

The disulfide linkages in this refolded and reoxidized protein were determined as for the pro-part from the mature enzyme. The separation by HPLC of the peptides obtained after the cleavage with trypsin led to an elution profile identical to that of the native residual pro-part, confirming that the native-state pattern of disulfide bonds had been formed. Furthermore, peptides with free cysteine were not observed, indicating that reoxidation was complete.

The far-UV CD spectrum of the refolded and reoxidized residual pro-part shows the same, mostly β-sheet, conformation as that folded from the oxidized state. The fluorescence and the near-UV CD spectra (Figure 2), which are identical to those of the native form, confirm that the residual pro-part can be oxidatively refolded in vitro from the reduced state.

The residual pro-part retains the nativelike conformation in the absence of denaturant even if both disulfide bonds are reduced (see preceding section). To determine whether the formation of disulfide bonds is essential for conformational folding to take place, fully reduced and denatured residual pro-part was diluted into refolding buffer containing 20 mM DTT. After 2 min, the fluorescence spectra in 0.2, 0.8, and 1.5 M GdmCl were scanned and compared to those measured at the same concentrations of GdmCl and DTT added to the folded, disulfide-bonded residual pro-part. The fluorescence intensities and maxima of the residual pro-part prepared in either of the two ways were the same, indicating that the reduced and denatured residual pro-part can rapidly adopt the nativelike conformation without prior formation of disulfide bonds.

Oxidative refolding was carried out in a significant concentration of denaturant. That this results in the same conformation as that of the residual pro-part isolated from cathepsin C is supported by the facts that (i) the denaturant concentration used was within the native plateau region of the oxidized and reduced pro-part, (ii) the fluorescence spectra of the reduced protein obtained by reduction of the folded oxidized protein, and by refolding the reduced and unfolded protein in 0.2 M GdmCl and in 1.5 M GdmCl (equivalent to 3 M urea + 0.6 M GdmCl), were superimposable, (iii) the disulfide linkage was the same and the sensitive near-UV CD spectra were indistinguishable, and (iv) the disulfides are solvent-exposed and can be reduced in the absence of denaturant which, together with the above experiments, indicates that the overall reductive unfolding process is fully reversible.

DISCUSSION

The Function of the Residual Pro-Part. The larger propeptides in proteinase zymogens have been proposed to have two main functions. First, they are responsible for suppressing the activity of the enzyme and, in certain cases, including the cysteine proteinase cathepsin L (*30*), are able to act directly as inhibitors. Second, the propeptides of α-lytic protease, subtilisin, and the cysteine proteinase from *Bombyx mori* are essential for enabling the respective enzymes to surmount the activation barrier between their partially folded and native states (*3*, *4*). It has generally been assumed, by analogy, that the propeptides of the papain-like cysteine proteinases assist folding.

Like these enzymes, mature cathepsin C cannot be reversibly unfolded, neither activity nor the oligomeric structure being regained even when refolding was attempted in the presence of the already folded residual pro-part (data not shown). The precipitation of the 'heavy' and 'light' peptides of the catalytic moiety, on transferring to native solvent conditions, forms the basis of a simple procedure for isolating homogeneous residual pro-part from the recombinant enzyme. This was also possible using enzyme extracted from human tissue, but a substantial proportion of this residual pro-part is nicked during activation (*20*), markedly reducing the yield of intact residual pro-part. With respect to amount and sites of glycosylation, the pro-part of recombinant human cathepsin C is comparable to that isolated from the natural enzyme, and its stability and folding kinetics are shown to differ little, if at all. The higher apparent molecular mass values of 17 and 25 kDa determined by size exclusion chromatography and reducing SDS−PAGE, respectively, as well as the high solubility of the residual pro-part, are attributable to the high degree of glycosylation of the 13.5 kDa protein (*20*). In addition, the linkages of the two disulfide bonds present in the active enzyme have been identified. The recombinant material is therefore a suitable model for drawing conclusions about the folding of the natural human procathepsin C in vivo.

The residual pro-part isolated from the mature enzyme folds spontaneously and reversibly to a stable (7.7 kcal·mol^{-1}) and cooperative globular structure. Tryptophans well shielded from the solvent, the presence of well-defined secondary structure and stable tertiary structure, and the absence of ANS binding confirm that the structure is a well packed and compact globule.

Disulfide bonds can make a substantial contribution to the stability of a protein. Proteins such as bovine pancreatic trypsin inhibitor, ribonuclease A, and α-lactalbumin adopt the random coil conformation after full reduction, even under physiological conditions and in the absence of denaturant [reviewed in (*31*)]. However, many proteins survive in a quasi-native conformation after reduction of one or two disulfides, including BPTI (*32*), RNaseT1 (*33*, *34*), and RnaseA (*35*), while TEM-1 β-lactamase (*36*) and *Erythrena caffra* trypsin inhibitor (*37*) even retain their function when the disulfide bonds are reduced, the only effect in these cases being to lower the stability. The two disulfide bonds of the residual pro-part, which can be reduced in the folded state, stabilize the native conformation by only 2 kcal·mol^{-1}. Analysis of the cDNA sequence (*18*) shows that the three glycosylated asparagines in the residual pro-part (*20*) precede or follow three of the four cysteine residues. The environment of the cysteines thus accounts for the susceptibility of the disulfides to reduction in the folded state, and indicates that both disulfides are exposed to solvent (*38*). The reduced form thus exhibits a quasi-native conformation, retaining the secondary structure, the dynamics, and most or all of the persistent tertiary interactions of the oxidized form. The folded state with reduced disulfide bonds has the same spectroscopic properties, whether reached from the unfolded, reduced state or by the reduction of disulfide bonds in the folded state. The residual pro-part folded in the absence of disulfide bonds is thus a thermodynamically stable intermediate with nativelike tertiary structure. The folded state is achieved, in a manual mixing experiment in 1.48 M GdmCl,

in less than 30 s, the greater proportion of the fluorescence change, as for the oxidized form, occurring much faster. The dependence of folding rate constants on denaturant concentration observed for the oxidized form (Figure 5) shows that, in the absence of GdmCl, it will fold significantly faster. These results show that the residual pro-part can fold rapidly relative to the rate of biosynthesis of up to 20 residues per second (*39*).

The two disulfide bonds formed on refolding and reoxidation have been shown to be the same as those in the residual pro-part folded in vivo during synthesis of the proenzyme. The overall reversibility of oxidative refolding is strong evidence that the conformation of the spontaneously folding pro-part determined here in vitro is that which this domain can achieve when folding off the ribosome in vivo. The kinetics of folding and the exposure of the disulfides to reduction show that folding and oxidation are sequential and that this will be the likely mechanism in vivo. The oxidation of solvent-exposed cysteines in a nativelike intermediate is usually relatively fast, as in the case of intermediate (30−51,5−55) of bovine pancreatic trypsin inhibitor (*40*) and in vivo is even faster, being catalyzed by protein disulfide isomerase. In the case of the residual pro-part, even if the rate of disulfide bond formation were slow as compared to protein synthesis, the reduced residual pro-part with nativelike conformation would still form a matrix which could interact noncovalently with the rest of the enzyme.

Concentrations in excess of 2 M GdmCl are needed to initiate cooperative denaturation of cathepsin C (*28*). This, coupled with a C_m for the residual pro-part of 3.7 M GdmCl and its high thermodynamic stability, indicates that the residual pro-part stabilizes the mature enzyme. The residual pro-part thus meets the double criteria of being able to assist folding of the catalytic moiety in vivo by folding rapidly to a native or nativelike conformation, and by being able to interact with and stabilize the catalytic moiety of the proenzyme as it folds. These results also confirm that the folding of the proenzyme is sequential.

The Folding of the Oxidized Residual Pro-Part. The residual pro-part with disulfide bonds intact folds rapidly, faster than the reduced form. The rapid acquisition of 70% of native ellipticity (within 7 s in these experiments) is characteristic of many proteins (*41*). The stopped-flow burst phase entails 85% of the total change in fluorescence but, interestingly, only 30% of the increased seclusion from solvent of tryptophan that takes place on folding (Table 2). The major part, at least, of the change in intrinsic fluorescence must therefore be ascribed to specific interactions of tryptophans within a loose conformation, rather than to their sequestration from solvent accompanying nonspecific, hydrophobically driven collapse. This is consistent with a localized region containing nativelike aromatic interactions, an initiation site (*42*), the remainder of the structure being relatively poorly packed, as shown for barnase. This would be in keeping with the limited transient ANS binding.

The preparation of residual pro-part from cathepsin C isolated from human tissue comprises approximately 50% intact pro-part, the remainder being cleaved at one of two positions into two chains, shown here to be joined by disulfide bonds 30−118 and 54−136. It exhibits a well-defined tertiary structure by near-UV CD, similar to that of the recombinant pro-part (results not shown), despite having been unfolded in the procedure for dissociation from the catalytic moiety (see Experimental Procedures). The unfolding transition (Figure 4A and Table 1) and the kinetics of folding and unfolding (Table 3) show that the cleaved form is capable of folding, at least in the presence of the disulfide bond. This suggests that the residual pro-part comprises two folding domains of 10 and 3.5 kDa. By analogy to a multisubunit protein, whose subunits may collapse and associate to a molten globule conformation (*43*), the kinetics of folding of the intact residual pro-part may be interpreted in terms of rapid collapse, association, and mutual stabilization of the two folding units, followed by slower conformational rearrangement.

The unfolding kinetics of the intact residual pro-part are unusual in exhibiting more than one phase by intrinsic fluorescence, fluorescence quenching, and far-UV ellipticity (Table 2 and Figure 7A). Most proteins unfold with a simple, kinetically two-state transition, without the accumulation of intermediates. The multiphasic nature of the unfolding of the residual pro-part raises the question as to whether the folded state is homogeneous, particularly in light of reports that the complex unfolding of dihydrofolate reductase (*44*) and staphylococcal nuclease (*45*) results from the multiplicity of the native state. The sharp equilibrium transition and high thermodynamic stability of the folded state of the residual pro-part (Figure 4A, Table 1), however, make it unlikely that species with relatively large differences in free energy are present. The alternative possibility, that intermediates accumulate during unfolding, is not well understood (*46*, *47*), although some have been ascribed to proline isomerization during unfolding (*48*). The multiphasic folding observed here may be related to the proposed presence of two folding units.

In conclusion, the excised part of the propeptide of cathepsin C (the activation peptide) is similar in size and has some sequence similarity to the complete propeptide of cathepsin L, which is known to fold independently (*49*). The complete propeptides of cysteine proteinases are known in certain cases to assist folding of the catalytic region when attached covalently (*14*) and when separate (*16*). The activation peptide of cathepsin C may therefore assist folding.

The folded residual pro-part has been shown here to interact strongly with the catalytic moiety of cathepsin C, stabilizing it in its proteolytically active form. The ability of the residual pro-part to fold rapidly in vitro to a nativelike form means that, during synthesis of the proenzyme in vivo, it would be already folded and able to interact with the catalytic moiety as soon as the latter can present an approximately complementary surface. It appears therefore that, whatever the role of the activation peptide in folding, the residual pro-part assists folding of the proenzyme in vivo by stabilizing a late intermediate.

Since cathepsin C is the only known tetrameric cysteine proteinase, it is also possible that the additional, sequentially unique residual pro-part domain is involved in dimerization of the proenzyme (S. W. Dahl, personal communication) and tetramerization of the mature enzyme (*23*). In this case, the stable fold of the residual pro-part domain would provide a surface for linking to other catalytic units or for self-association.

ACKNOWLEDGMENT

We are grateful to Dr. Igor Križaj for N-terminal analyses and to Prof. Vito Turk for his support.

REFERENCES

1. Silen, J. L., and Agard, D. A. (1989) *Nature 341*, 462−464.
2. Ikemura, H., Takagi, H., and Inouye, M. (1987) *J. Biol. Chem. 262*, 7859−7864.
3. Baker, D., Sohl, J. L., and Agard, D. A. (1992) *Nature 356*, 263−265.
4. Eder, J., Rheinnecker, M., and Fersht, A. R. (1993) *Biochemistry 32*, 18−26.
5. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) *Nature 395*, 817−819.
6. Anderson, D. E., Peters, R. J., Wilk, B., and Agard, D. A. (1999) *Biochemistry 38*, 4726−4735.
7. Bergman, L. W., and Kuehl, W. M. (1979) *J. Biol. Chem. 254*, 5690−5694.
8. Kolb, V. A., Makeyev, E. V., and Spirin, A. S. (1994) *EMBO J. 13*, 3631−3637.
9. Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F. U. (1999) *Nat. Struct. Biol. 6*, 697−705.
10. Danner, M., and Seckler, R. (1993) *Protein Sci. 2*, 1869−1881.
11. Seckler, R. (1997) in *Molecular Chaperones in the Life Cycle of Proteins* (Fink, A. L., and Goto, Y., Eds.) pp 391−413, Marcel Dekker, New York.
12. Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M., and Turk, V. (1996) *FEBS Lett. 384*, 211−214.
13. Smith, S. M., and Gottesman, M. M. (1989) *J. Biol. Chem. 264*, 20487−20495.
14. Ogino, T., Kaji, T., Kawabata, M., Satoh, K., Tomoo, K., Ishida, T., Yamazaki, H., Ishidoh, K., and Kominami, E. (1999) *J. Biochem. (Tokyo) 126*, 78−83.
15. Hou, W. S., Bromme, D., Zhao, Y., Mehler, E., Dushey, C., Weinstein, H., Miranda, C. S., Fraga, C., Greig, F., Carey, J., Rimoin, D. L., Desnick, R. J., and Gelb, B. D. (1999) *J. Clin. Invest. 103*, 731−738.
16. Yamamoto, Y., Watabe, S., Kageyama, T., and Takahashi, S. Y. (1999) *Arch. Insect. Biochem. Physiol. 42*, 167−178.
17. Nagler, D. K., Sulea, T., and Menard, R. (1999) *Biochem. Biophys. Res. Commun. 257*, 313−318.
18. Pariš, A., Štrukelj, B., Pungerčar, J., Renko, M., Dolenc, I., and Turk, V. (1995) *FEBS Lett. 369*, 326−330.
19. Hola-Jamriska, L., Tort, J. F., Dalton, J. P., Day, S. R., Fan, J., Aaskov, J., and Brindley, P. J. (1998), *Eur. J. Biochem. 255*, 527−534.
20. Cigić, B., Križaj, I., Kralj, B., Turk, V., and Pain, R. H. (1998) *Biochim. Biophys. Acta 1382*, 143−150.
21. Dolenc, I., Turk, B., Pungerčič, G., Ritonja, A., and Turk, V. (1995) *J. Biol. Chem. 270*, 21626−21631.
22. Nikawa, T., Towatari, T., and Katunuma, N. (1992) *Eur. J. Biochem. 204*, 381−393.
23. Muno, D., Ishidoh, K., Ueno, T., and Kominami, E. (1993) *Arch. Biochem. Biophys. 306*, 103−110.
24. Vanhove, M., Lejeune, A., and Pain, R. H. (1998) *Cell. Mol. Life Sci. 54*, 372−377.
25. Greene, R. F., and Pace, N. C. (1974) *J. Biol. Chem. 249*, 5338−5393.
26. Ackerman, R. J., and Robyt, J. F. (1972) *Anal. Biochem. 50*, 656−659.
27. Ryle, A. P., Sanger, F., Smith, L. F., and Kitai, R. (1955) *Biochem. J. 60*, 541−556.
28. Cigić, B., and Pain, R. H. (1999) *Biochem. Biophys. Res. Commun. 258*, 6−10.
29. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Žerovnik, E., and Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20−24.
30. Carmona, E., Dufour, E., Plouffe, C., Takebe, S., Mason, P., Mort, J. S., and Menard, R. (1996) *Biochemistry 35*, 8149−8157.
31. Creighton, T. E. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 301−351, W. H. Freeman and Company, New York.
32. Kortemme, T., Hollecker, M., Kemmink, J., and Creighton, T. E. (1996) *J. Mol. Biol. 257*, 188.
33. Pace, C. N., and Creighton, T. E. (1986) *J. Mol. Biol. 188*, 477.
34. Mayr, L. M., Willbold, D., Landt, O., and Schmid, F. X. (1994) *Protein Sci. 3*, 227.
35. Pearson, M. A., Karplus, P. A., Dodge, R. W., Laity, J. H., and Scheraga, H. A. (1998) *Protein Sci. 7*, 1255.
36. Vanhove, M., Guillaume, G., Ledent, P., Richards, J. H., Pain, R. H., and Frere, J.-M. (1997) *Biochem. J. 321*, 413−417.
37. Lehle, K., Wrba, A., and Jaenicke, R. (1994) *J. Mol. Biol. 239*, 276−284.
38. Holst, B., Bruun, A. W., Kielland-Brandt, M. C., and Winther, J. R. (1996) *EMBO J. 15*, 3538−3546.
39. Hershey, J. W. (1991) *Annu. Rev. Biochem. 60*, 717−755.
40. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497−526.
41. Kuwajima, K., Semisotnov, G. V., Finkelstein, A. V., Sugai, S., and Ptitsyn, O. B. (1993) *FEBS Lett. 334*, 265−268.
42. Freund, S. M., Wong, K. B., and Fersht, A. R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10600−10603.
43. Hlodan, R., and Pain, R. H. (1995) *Eur. J. Biochem. 231*, 381−387.
44. Hoeltzli, S. D., and Frieden, C. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9318−9322.
45. Kuwajima, K., Okayama, N., Yamamoto, K., Ishihara, T., and Sugai, S. (1991) *FEBS Lett. 290*, 135−138.
46. Bhuyan, A. K., and Udgaonkar, J. B. (1998) *Biochemistry 37*, 9147−9155.
47. Zaidi, F. N., Nath, U., and Udgaonkar, J. B. (1997) *Nat. Struct. Biol. 4*, 1016−1024.
48. Juminaga, D., Wedemeyer, W. J., Garduno-Juarez, R., McDonald, M. A., and Scheraga, H. A. (1997) *Biochemistry 36*, 10131−10144.
49. Jerala, R., Žerovnik, E., Kidrič, J., and Turk, V. (1998) *J. Biol. Chem. 273*, 11498−11504.

BI0008837