Folding-related Dimerization of Human Cystatin C*

Irena Ekiel‡§ and Magnus Abrahamson¶

From the ‡Pharmaceutical Biotechnology Sector, Biotechnology Research Institute, National Research Council of Canada, Montreal, Quebec H4P 2R2, Canada and the ¶Department of Clinical Chemistry, University of Lund, University Hospital, S-221 85 Lund, Sweden

With the aim to improve our understanding of the structural basis for protein self-association and aggregation, in particular in relationship to protein refolding and amyloid formation, folding-related processes for human cystatin C have been studied. Using NMR spectroscopy together with chromatographic and electrophoretic methods, a self-association process resulting in dimer formation for protein samples treated with denaturing agents as well as for samples subjected to low pH or high temperature conditions could be studied with amino acid resolution. In all three cases, the dimerization involves properly folded molecules and proceeds via the reactive site of the inhibitor, which leads to complete loss of its biological activity. This dimerization process has potential relevance for amyloid formation by the brain hemorrhage-causing Leu[68]-Gln variant of cystatin C. The results also indicate that cystatin C dimerization and inactivation may occur in acidified compartments *in vivo*, which could be relevant for the physiological regulation of cysteine proteinase activity.

The self-association and aggregation of proteins constitutes one of the least understood problems in protein chemistry.[1] Although some insight has been gained in the kinetic aspects of the aggregation process and in the field of its general theory (Glatz, 1992; Honig and Yang, 1995; De Young *et al.*, 1993), very little is known about its structural aspects and about the precise nature of the interactions involved. It is generally assumed that aggregation proceeds via a partially exposed hydrophobic core in a molten globule, but the very general character of the molten globule concept seriously limits any practical consequences. For controlled refolding of recombinant proteins and for stabilization of pharmaceutical formulations, as well as for the development of new therapies for aggregation-related disease states including amyloidosis, more development in this research area clearly is needed.

To add to our understanding of the molecular aspects of amyloid formation, in particular of the specific intermolecular interactions leading to fibril formation, we have started systematic studies of the properties of amyloidogenic proteins. Human cystatin C provides a good starting point for such studies. This small size inhibitor of cysteine proteinases is present in all human body fluids at physiologically relevant concentrations, being most abundant in the cerebrospinal fluid and in seminal plasma (Abrahamson *et al.*, 1986). Although cystatin C in its wild-type form has not been reported to form amyloid *in vivo*, its Leu[68] → Gln mutated variant (L68Q-cystatin C) is responsible for the dominantly inherited disorder called hereditary cystatin C amyloid angiopathy (Ghiso *et al.*, 1986; Jensson *et al.*, 1987; Palsdottir *et al.*, 1988). Expression systems for both wild-type and L68Q-cystatin C have been developed (Dalbøge *et al.*, 1989; Abrahamson and Grubb, 1994), and the three-dimensional structure of the human cystatin C-like type II cystatin, chicken cystatin, is well characterized by both x-ray crystallography and NMR spectroscopy (Bode *et al.*, 1988; Engh *et al.*, 1993). The conformation of the 120-residue polypeptide chain of cystatin C is very similar to that of chicken cystatin,[2] with a proteinase-interacting wedge-shaped side involving residues Arg[8]-Gly[11] of the N-terminal segment and two loop-forming segments constituting polypeptide turns in the main, β-pleated sheet structure of the molecule (segments Gln[55]-Gly[59] and around the single tryptophan residue, Trp[106]). An additional advantage for the study of cystatin C in the aspect of protein aggregation and amyloid formation is the fact that it normally exists in a monomeric form, in sharp contrast to many other amyloid-forming proteins which are found in multimeric, often heterogeneous, and condition-dependent forms (Brange *et al.*, 1987; Brader and Dunn, 1991; Colon and Kelly, 1992).

Only about 15 of the human body fluid proteins are proven to be amyloidogenic (Stone, 1990; Sipe, 1992). The formation of amyloid is often either directly related to specific point mutations or is enhanced significantly by point mutations (Frangione, 1991; Glenner *et al.*, 1991; Sipe, 1992). As is particularly clear in the case of transthyretin, many different point mutations which do not form a clear pattern can lead to amyloid formation. Therefore, there is a tendency to explain amyloidogenity as being caused by reduced stability of the proteins (Hurle *et al.*, 1994; McCutchen *et al.*, 1993). Earlier results (Abrahamson and Grubb, 1994) suggest that cystatin C and its L68Q variant can follow a similar pattern. In the present investigation, we have undertaken a detailed study of wild-type cystatin C under conditions leading to its folding-related self-association, with the aim to define the molecular background to the events leading to amyloid formation and physiological inactivation of L68Q-cystatin C.

§ To whom correspondence should be addressed: Pharmaceutical Biotechnology Sector, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Ave., Montreal, Quebec H4P 2R2, Canada. Tel. 514-496-6329; Fax: 514-496-5143.

[1] Throughout this paper, definitions of "self-association" as a reversible process involving the interactions of two or more native protein molecules (with reversible precipitation of the protein as a possible consequence) and "aggregation" as the interaction of two or more denatured protein molecules (which often leads to practically irreversible precipitation) are used, as suggested by Cleland *et al.* (1993).

[2] I. Ekiel, unpublished NMR observations.

EXPERIMENTAL PROCEDURES

Materials—Recombinant cystatin C, with structure and functional properties identical with those of the native inhibitor isolated from human urine, was produced in *E. coli* and purified using previously described methods (Abrahamson *et al.*, 1988; Dalbøge *et al.*, 1989). Superdex™ 75 (preparative grade) resin was obtained from Pharmacia LKB, Uppsala, Sweden. Guanidine hydrochloride (GdnHCl)[3] was purchased from BDH Chemicals Ltd (Aristar quality), and Z-Gly-*p*-nitrophenyl ester from Armand-Frappier, Laval, Quebec, Canada. All other reagents were purchased from Sigma and used without further purifications. Affinity-purified papain was obtained as a gift from Dr. A. C. Storer and R. Dupras.

Preparation of Self-associated Forms of Cystatin C—Self-associated forms of cystatin were studied by incubation of 1 mg/ml (75 μM) solutions of cystatin C in 50 mM sodium phosphate buffer, pH 6.0, for GdnHCl denaturation experiments, and in 50 mM sodium phosphate buffer, pH 6.7, containing 0.1 M NaCl, for temperature-dependent experiments. Experiments at low pH were performed in 0.1 M sodium acetate buffer or acetic acid solutions, which were prepared in D_2O. Samples were incubated for 30 min in temperature experiments and 11–17 h for the pH and GdnHCl experiments. Such treated samples were used directly for NMR experiments as well as for agarose gel electrophoresis. For comparisons of self-associated forms formed after different treatments, GdnHCl solutions or low pH solutions were diluted quickly into a standard phosphate buffer (50 mM phosphate, pH 6.7, containing 0.1 M NaCl), in which no cystatin C self-association was detected even after 3 months of incubation. No conversion of dimers to monomers was observed during storage of diluted samples. Samples incubated in GdnHCl were diluted to a final GdnHCl concentration not higher than 0.1 M and concentrated back to their original volumes using N_2 pressure filtration in stirred cells equipped with YM3 or YM10 membranes (Amicon) or in centrifugal microconcentrators (Centricon 3 or 10; Amicon). The buffer was then further exchanged at least 50-fold.

NMR Spectroscopy—Nuclear Magnetic Resonance (NMR) spectra were collected at 500 MHz using Bruker AMX500 and AM500 spectrometers; 10% (v/v) D_2O was added for locking purposes in H_2O samples. For samples in heavy water, pD values are given as pH-meter readings at 23 °C without any further corrections. Chemical shift values are given in parts per million relative to the internal 2,2-dimethyl-2-silapentane-5-sulfonic acid. NMR signals were identified using the assignments performed for the isotope-labeled (^{13}C and ^{15}N) cystatin C[4] under optimal conditions (1–2 mM solutions at pH 6.0, at 30 °C). These assignments were extended to the conditions needed in this work by titration-type experiments between the standard conditions and those used in a particular experiment. For most informative high temperature NMR experiments, to verify the assignments, two-dimensional TOCSY experiments with WALTZ-17 or MLEV mixing sequence (Braunschweiler and Ernst, 1983; Bax and Davis, 1985) were performed at 40 °C, the highest temperature at which two-dimensional experiments could be performed at 1.5 mM concentrations without samples undergoing dimerization. The mixing time ranged from 40 to 114 ms and 360 free induction decays of 1024 complex points each (120 scans per free induction decay) were acquired. Solvent suppression was accomplished using presaturation. Aromatic ring signals were assigned using NOESY spectra as described elsewhere.[4] In this way, signal assignments in one-dimensional spectra were done for representative amino acids from all parts of the molecule. Most importantly, a good set of assigned signals was obtained for amino acids in the vicinity of the active site (Trp[106], Val[104], Tyr[102], and Ala[58]).

NMR experiments were run in both regular and heavy water. From the NH signals observed in water, NH of the Trp[106] residue is particularly important (see below). In heavy water, as NH signals are exchanged to ND, many aromatic signals can be interpreted clearly. To give an overview of advantages given by both methods and to avoid repetitions, results for the GdnHCl experiments will be presented only for the H_2O solutions, and results for temperature and pH experiments only for the D_2O solutions.

Other Methods—After long incubations at room temperature, cystatin C samples were checked for potential degradation using mass spectroscopy (API III MS/MS System, Sciex, Thornhill, Ontario, Canada). Agarose gel electrophoresis was performed according to Jeppsson *et al.*
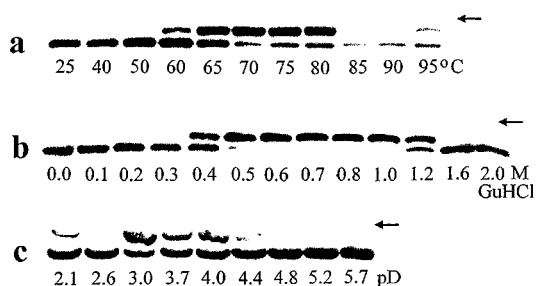
---

[3] The abbreviations used are: GdnHCl, guanidine hydrochloride; SEC, size exclusion chromatography; PAGE, polyacrylamide gel electrophoresis; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy.

[4] I. Ekiel and M. Abrahamson, manuscript in preparation.

FIG. 1. **Agarose gel electrophoresis under partially unfolding conditions.** Isolated recombinant cystatin C was incubated at 75 mM concentration. *a*, for 30 min at various temperatures in 50 mM sodium phosphate buffer, pH 6.7, containing 0.1 M NaCl; *b*, for 17 h in varying concentrations of GdnHCl (*GuHCl*) in 50 mM sodium phosphate buffer, pH 6.0, at room temperature; *c*, for 11 h at various pD values. The incubation mixtures were analyzed by electrophoresis in 1% (w/v) agarose gels. The direction of the electric field is indicated to the *left*. The points of sample application are marked by *arrows*.

(1979). Nondenaturing PAGE was carried out in continuous 16.5% acrylamide gels using the pH 4.0 alanine/acetate buffer system described by Jovin (1973). Size exclusion chromatography (SEC) was performed using a Superdex™75 column (30 × 1 cm, flow rate 38 cm/h) equilibrated in 50 mM sodium phosphate buffer, pH 6.7, containing 0.1 M NaCl. Ribonuclease A (M_r = 13,700), chymotrypsinogen A (M_r = 25,000), ovalbumin (M_r = 43,000), and bovine serum albumin (M_r = 57,000) (low molecular weight get filtration kit from Pharmacia) were used for calibration of the column. Inhibitory activity of cystatin C samples was determined by titration of 0.1 μM papain solutions using 100 mM Z-Gly-*p*-nitrophenyl ester as a substrate, as described before (Berti, 1993).

RESULTS

Cystatin C Self-association under Mildly Denaturing Conditions—From the results of preliminary SEC experiments to find conditions leading to self-association or aggregation of wild-type cystatin C, a self-association or aggregation process was apparent for protein samples treated with denaturing agents, as well as for samples subjected to low pH or high temperature conditions. Moreover, it was noted that the process observed was very slow. It was, therefore, possible to follow it as a function of time using NMR spectroscopy and to trap associated forms by a quick transfer to standard conditions. It was, furthermore, demonstrated by both SEC and NMR spectroscopy that stationary and trapped cystatin C associated forms were very similar, if not identical (data not shown). Under close to unfolding conditions, the cystatin C associates observed were in an equilibrium with monomers (possibly unfolded), and trapping lead to some back conversion to monomers. However, this conversion was sufficiently slow to allow samples under partially denaturing conditions to be analyzed by agarose gel electrophoresis in nondenaturing buffer.

Screening of Conditions Leading to Cystatin C Self-association—Agarose gel electrophoresis was selected as a method for screening of conditions under which cystatin C self-association occurs, as it has been seen earlier in a study of the Leu[68] → Gln mutated cystatin C (L68Q-cystatin C) that this method is useful for simultaneously detecting dimer formation (as a distinct band shift) and higher order aggregates leading to precipitation (Abrahamson and Grubb, 1994). In Fig. 1, results are shown demonstrating wild-type cystatin C self-association induced by high temperature (*a*), by guanidine hydrochloride (*b*), and by low pH (*c*). In all three experiments, a region of extensive self-association was observed, with a profile resembling that of the frequently seen aggregation of partially unfolded proteins (Ghélis and Yon, 1982; Mitraki and King, 1989). The main cystatin C self-association product was small, homogeneous, and soluble at 75 μM concentration according to the agarose gels. Extensive precipitation was observed only in the
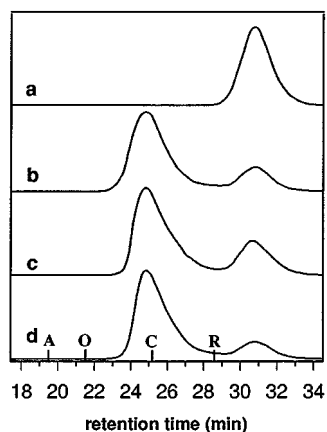
FIG. 2. **Size exclusion chromatography of cystatin C dimers.** Dimer-containing cystatin C samples obtained under conditions described in the legend to Fig. 1 were trapped by rapid transfer to a 50 mM phosphate buffer, pH 6.7, containing 0.1 M NaCl. *a*, monomeric cystatin C; *b*, dimer produced by an incubation in 0.8 M GdnHCl for 16 h at room temperature; *c*, dimer produced at pD 3.0 (16-h incubation at room temperature); *d*, dimer formed by incubation at 68 °C for 30 min. Retention times for the calibration standards bovine serum albumin (*A*), ovalbumin (*O*), chymotrypsinogen (*C*), and ribonuclease A (*R*) are indicated.

temperature dependence experiment above 80 °C, where cystatin C unfolds, as was evident from NMR spectra run at high temperatures (data not shown). The small self-associated form seemed to have reduced solubility compared with the monomeric molecule, as further aggregation and precipitation were observed in repeated experiments at cystatin C concentrations above 0.4 mM (not shown).

In the experiments with increasing concentrations of guanidine hydrochloride, self-associated cystatin C started to appear at 0.3–0.4 M GdnHCl, and the amount of trapped self-associated form started decreasing at 1.2 M GdnHCl. Similar results were obtained at low pH values, with maximum amounts of self-associated cystatin C observed in the pH range 3.0 to 4.4. The relative amount of self-associated cystatin decreased rapidly when pH was lowered further (Fig. 1*c*).

Indications for Formed Cystatin C Self-associated Species Being Dimers by SEC—In parallel to electrophoretic analysis of cystatin C samples under different conditions, SEC experiments were performed on selected samples. As seen from the chromatograms in Fig. 2, samples indicated to contain partially self-associated cystatin C by agarose gel electrophoresis, from either the temperature, pH, or GdnHCl experiments, all gave two peaks in SEC, with retention times of 25.0 min (self-associated form) and 30.7 min (monomeric cystatin). The retention time of the first peak was very similar to that of chymotrypsinogen ($M_r = 25,000$) and, therefore, likely represented a cystatin dimer. No soluble species larger than dimers could be detected. However, some insoluble, precipitated material was produced, especially in samples incubated at high temperatures. In the SEC experiments, it was observed that monomeric cystatin C was more retained than the similarly sized protein, ribonuclease A. This excessive retention of cystatin C is caused by interactions with the matrix, most likely via the inhibitor's proteinase binding site, as a cystatin C variant with the single amino acid substitution Trp[106] → Gly (W106G-cystatin C) (Hall *et al.*, 1995) gave a retention time of 28.4 min, *i.e.* quite different from that for wild-type cystatin C, but similar to that for ribonuclease A (28.7 min). The wild-type cystatin C-matrix interaction clearly changes in the dimeric state, as the retention times for wild-type and W106G-cystatin C dimers were identical. The conclusion that the self-associated species studied were dimers was additionally supported by nondenaturing
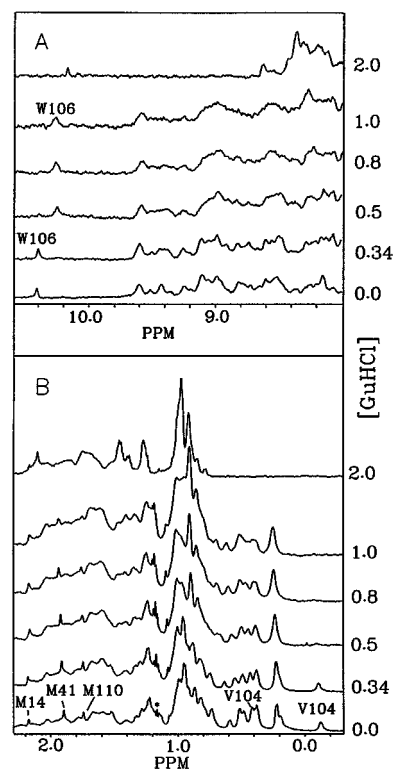

FIG. 3. **¹H NMR spectra of cystatin C in various concentrations of GdnHCl.** Spectra were run in 50 mM phosphate buffer, pH 6.0, containing between 0.0 and 2.0 M GdnHCl at 30 °C after overnight (17 h) incubation. *A*, downfield region; *B*, upfield region.

PAGE (data not shown).

NMR Data for Cystatin C Dimerization Induced by GdnHCl—NMR spectra of cystatin C (75 μM) were run as a function of the GdnHCl concentration (Fig. 3). It is clear from these spectra that the dispersion of signals characteristic for folded proteins is preserved up to 1.0 M GdnHCl. In particular, a group of upfield shifted methyl group signals indicate a retained tertiary structure. Also, three methyl group signals of methionine residues retain the chemical shifts of a properly folded, monomeric protein and collapse only at higher concentrations of between 1 and 2 M GdnHCl. In the NH region, downfield shifted signals, typical for a β-sheet structure, are preserved up to 1 M GdnHCl, but change dramatically between 1 and 2 M GdnHCl, supporting unfolding of the protein in this concentration range of denaturant. The NMR results moreover demonstrate a clear transition region between 0.5 and 1 M GdnHCl, where protein conformation seems to be retained, but specific changes are observed for some amino acid residues, including Trp[106] and Val[104] (a more specific description of these and other NMR changes in relation to the conformational changes will be presented elsewhere). It is important to note that before the protein starts unfolding, the NH proton of Trp[106] shows two discrete signals, shifted by 0.16 ppm, indicating existence of a transition between two states. Comparison with Fig. 1 demonstrates a correlation between the formation of the cystatin C dimers and the presence of the characteristic signal shifts most clearly seen for Val[104] and Trp[106] (Fig. 3). Dimers as detected by agarose electrophoresis are not present in the denaturant concentration of 2.0 M, for which concentration NMR indicated unfolding of the protein. Additionally, in the NMR spectra, an overall broadening of the signals is evident in the 0.5–1.0 M GdnHCl concentration range, consistent with a self-association process taking place.

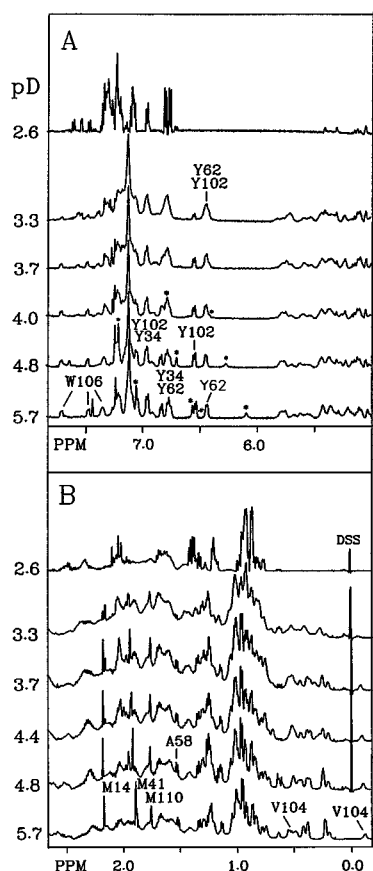*NMR Studies of pH-induced Cystatin C Dimerization*—NMR

FIG. 4. **Downfield (*A*) and upfield (*B*) regions of NMR spectra of cystatin C as a function of pD.** Spectra were run in 0.1 M sodium acetate buffer in heavy water. Titrating histidine signals are marked with *asterisks*.

spectra were run for cystatin C solutions in deuterated acetate buffer, in the pD range 2.6–5.7 (Fig. 4). As expected, more pronounced changes in chemical shifts were observed than in the GdnHCl experiment, as glutamic acid, aspartic acid, and histidine side chains undergo changes in the ionization state in the pD range covered in this experiment. However, important overall features of the spectra remained little changed in the pH range 5.7–3.3; there is a set of upfield shifted methyl group signals characteristic for tertiary structure, and the positions of methyl groups of all three methionine residues stay essentially the same with the exception of a small shift (0.02 ppm) for that of Met[14] (a similar shift was observed also in the GdnHCl-induced dimerization). Characteristically, downfield-shifted β-sheet α-protons remained unchanged in this pD range. The most clear interpretation of the individual signals is possible in the aromatic area, where most of the aromatic side chain signals retain their position in the pD range between 3.3 and 5.7 (with the exception of histidines titrating in this range). However, two aromatic side chains clearly differ in behavior. As can be seen most clearly by comparing spectra at pD 4.8 and 3.3, intensities of the Tyr[102] and Trp[106] signals for the monomeric cystatin decrease with decreasing pD and new signals appear in parallel with dimerization (*cf.* Fig. 1). For example, a decrease of the intensity of the Tyr[102] signal at 6.54 ppm is accompanied by an increase in the intensity of the broad signal at 6.42 ppm, where the Tyr[102] signal moves upon dimerization. At pD 3.3, similar changes as for the GdnHCl-induced dimerization appear in the methyl group signals of Val[104] (notice a decrease of the Val[104] signal at −0.1 ppm between pD 4.8 and 3.3); also the Ala[58] methyl group signal at 1.53 ppm changes. More dramatic changes in the NMR spectra are observed below

pD 3.3, where both upfield-shifted methyl group signals and downfield-shifted α-proton signals disappear. In addition, methionine methyl group signals dramatically change position, with nearly complete loss of the dispersion characteristic for them in folded proteins. In the aromatic region, the chemical shifts undergo large changes and approach those for unfolded proteins (Wishart and Sykes, 1994; Wüthrich, 1986). These changes in chemical shifts are accompanied by a decrease in linewidths (Fig. 4), which is typical for protein unfolding.

NMR Data for Temperature-induced Cystatin C Dimers—At elevated temperatures, NMR spectra of proteins generally exhibit motion-induced increased resolution. A more detailed analysis of cystatin C spectra at high temperatures was therefore carried out, as additional information about the nature of the self-associated forms could be gained. In order to study the dimerization events, special conditions were chosen to minimize higher aggregation (by using a low protein concentration of 75 μM and only a moderately high temperature of 62 °C). Under these conditions, dimerization was slow (hours), and, as an additional advantage, the important signals of branched amino acid methyl groups are well dispersed. To be able to interpret spectra under these conditions, assignments from low temperature NMR at 30 °C[4] were extrapolated by running temperature dependence of one-dimensional spectra (Fig. 5*B*) as well as two-dimensional TOCSY (Fig. 5*A*) and NOESY (not shown) spectra at 40 °C. In the one-dimensional spectra, 13 upfield shifted methyl groups from eight amino acids can be followed, as can methyl groups of the three methionine and some other alanine and threonine residues. All together, signals of 21 amino acids could be assigned easily at high temperature (see Fig. 6 and Table I). Although some of these signals undergo non-negligible shifts as a function of temperature, most of them were below 0.06 ppm in the aliphatic and even smaller in the aromatic region (spectra not shown). They were directly correlated to temperature changes and did not change with time; therefore, they could be attributed to small conformational changes and/or increasing motion in the molecule. At 62 °C, cystatin C clearly preserves its folded structure. After longer (overnight) incubation, dimerization occurs.

As can be seen in Fig. 6, *A* and *B*, NMR spectra of the monomeric and dimeric cystatin C are very similar. In particular, most of the characteristic signals which could be assigned easily and followed at high temperature were essentially not changed during dimerization. That includes amino acids in the β-sheet (Tyr[62], Leu[68], Leu[47], Met[110], and Phe[99]) as well as in the α-helix (Tyr[34] and Ala[30]). Also unchanged are Tyr[42] and Met[41] in the loop between the α-helix and the β-sheet, Leu[91] in the loop area of the β-sheet, as well as all histidines and most of the alanine methyl groups. The overall envelope, especially of the aromatic and α-protons of the β-sheet, also remains the same. That is in contrast to Trp[106], Tyr[102], and Val[104] signals, which undergo relatively large (above 0.2 ppm) changes. These latter changes develop slowly, on the time scale parallel to the formation of dimers (data not shown). Similarly, time-dependent changes were detected for the Ala[58] methyl group signal.

Comparison of Dimers Formed under Different Conditions—To verify that the cystatin C dimers formed under different conditions are identical, dimers were produced under optimal conditions for each of the three methods described above, and analyzed after buffer exchange. These dimer preparations gave identically migrating agarose gel electrophoresis bands and displayed identical retention times in SEC experiments (Fig. 2). The most detailed comparison was made using NMR spectroscopy. Fig. 6 (*spectra b*, *c*, and *d*) compare the upfield-shifted methyl group and aromatic regions of the proton NMR spectra for the three dimers. For the best comparison,
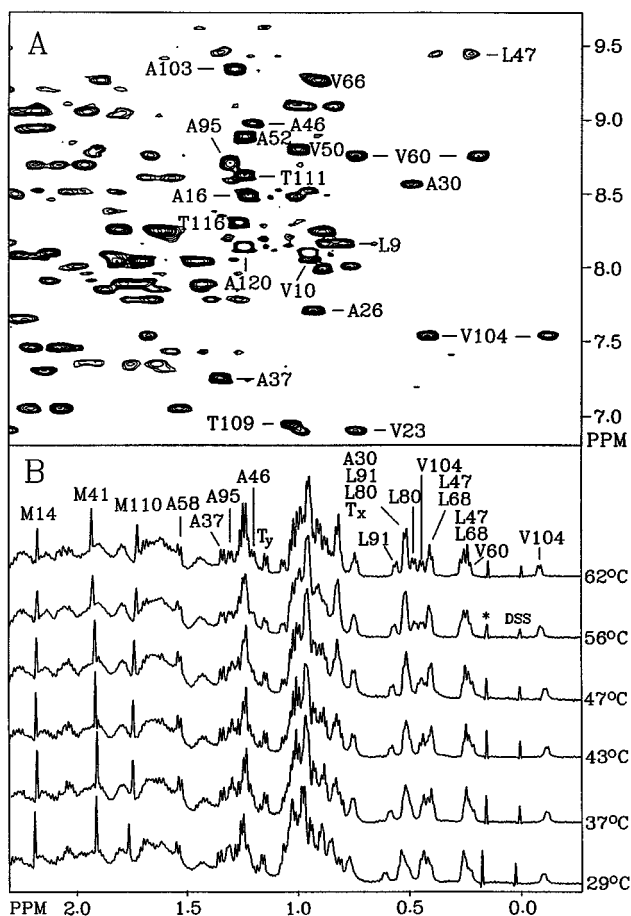
FIG. 5. **The upfield region of the NMR spectra of monomeric cystatin C.** *A*, fragment of the two-dimensional TOCSY spectrum at 40 °C (mixing time 114 ms) showing assignments for methyl group signals used as markers in one-dimensional spectra. *B*, temperature dependence of one-dimensional spectra. Spectra were run in sodium phosphate buffer containing 0.1 M NaCl at pD 6.7, with data collection time of 2 h (4000 accumulations). T_x and T_y denote two unassigned threonine methyl group signals. Signals of residues Leu[68], Leu[80], and Leu[91] were assigned using the aliphatic region of the same TOCSY spectrum, and the methyl group signals of three methionines using a two-dimensional NOESY spectrum (see text).

FIG. 6. **Comparison of the aromatic (A) and aliphatic (B) regions of ¹H NMR spectra of monomeric and trapped dimeric cystatin.** Spectra were run at 62 °C in 50 mM sodium phosphate buffer containing 0.1 M NaCl, at pD 6.7 in heavy water, using 4000 accumulations for each sample. *a*, monomeric cystatin C; *b*, dimer formed by an incubation at 62 °C for 32 h (notice temperature-induced exchange of histidine signals); *c*, dimer formed by an incubation for 16 h in 0.8 M GdnHCl solution; *d*, dimer formed at pD 3.0 for 16 h. Three histidine hydrogens which exchanged to deuterium during incubation at high temperature (*spectrum b*) are labeled with *asterisks*.

high temperature (62 °C) was selected for the optimal resolution and separation of signals in the aliphatic region (see Fig. 5). It was tested in an independent experiment (not shown) that at 62 °C during the time period for NMR data collection (2 h) only about 5% of the monomer was converted to dimer and that spectra of dimers practically do not change with time under these conditions. Fig. 6 shows a very close similarity of the spectra in both the aromatic and methyl group region, therefore supporting that all three differently formed dimers are structurally very similar. NMR spectra of three dimeric samples were also very similar at lower temperatures (between 30 and 60 °C), providing an additional support for the similarities of structures.

Dimer Formation Leads to a Loss of Cystatin C Inhibitory Activity against Cysteine Proteinases—X-ray crystallography results for the avian homologue to human cystatin C, chicken cystatin, as well as for the more distantly related human inhibitor, stefin B (also called cystatin B) in complex with papain (Bode *et al.*, 1988; Stubbs *et al.*, 1990) indicate that cystatins generally interact with cysteine proteinases via a contact area involving a very hydrophobic surface built from two loop-forming segments (comprising amino acids QIVAGVN and YAVP-WQGT, at positions 55–61 and 102–109, respectively, for hu-
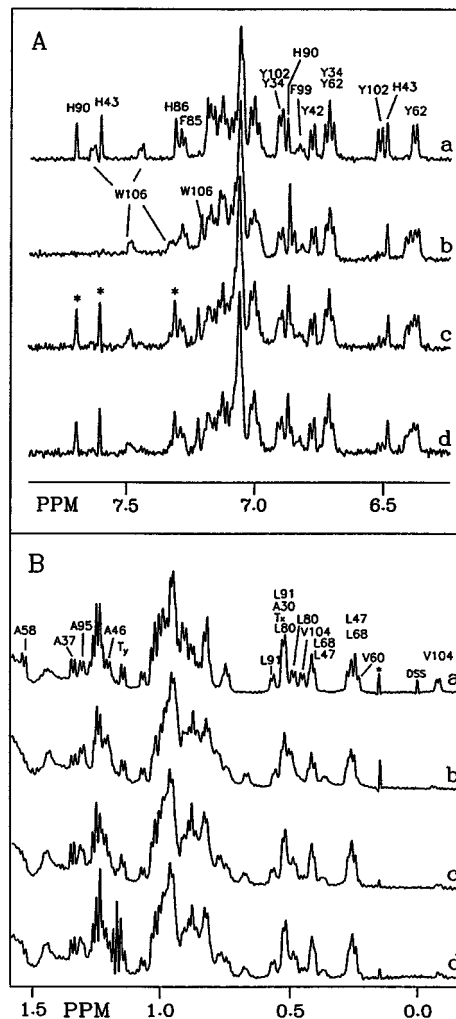
man cystatin C). The NMR results shown above for cystatin C residues in these two segments clearly indicate either some conformational changes in the contact region of the molecule or direct interactions via the proteinase-interacting hydrophobic surface. Therefore, it seemed likely that upon dimerization cystatin C would lose its activity as a cysteine proteinase inhibitor. To examine this possibility, dimers were trapped by buffer exchange and purified using size exclusion chromatography, and their activity was assayed against papain (an enzyme with K_i for the interaction with monomeric cystatin C in the picomolar range). Under conditions resulting in a mixture of monomeric and dimeric cystatin C upon trapping, the papain inhibitory activity of the mixture was proportional to the amount of the monomeric form of cystatin C. The monomeric form purified by SEC from such mixtures was in every aspect indistinguishable from freshly isolated (monomeric) cystatin C (NMR spectra, R_F values at SEC, full activity against papain). Incubation times of the papain-cystatin C dimer mixture were varied for up to 1 h, with no signs of inhibition developing. From this result it was concluded that the cystatin C dimer-

TABLE I
*Chemical shifts (in parts per million) for those amino acids in cystatin C, signals of which
were used in the interpretation of one-dimensional spectra*

Data were acquired at 40 °C, chemical shifts were referenced to internal 2,2-dimethyl-2-silapentane-5-sulfonic acid.

Residue	NH	Hα	Hβ	Others
Ala[30]	8.49	3.18	0.51	
Tyr[34]	8.70	4.03	3.28, 3.08	Hδ 6.96; Hε 6.77
Ala[37]	7.21	4.33	1.35	
Tyr[42]	8.11	4.80	2.79	6.82; 7.24
His[43]	8.79	4.46	2.77, 2.59	Hδ$_2$ 6.53; Hε$_1$ 7.71
Ala[46]	8.94	4.36	1.21	
Leu[47]	9.40	4.28	1.36, 1.21	Hγ 1.17; Hδ 0.41, 0.26
Ala[58]	8.20	4.72	1.53	
Val[60]	8.71	4.68	1.67	Hγ 0.76; 0.23
Tyr[62]	9.57	5.09	2.74, 2.71	Hδ 6.75; Hε 6.42
Leu[68]	9.58	4.92	1.44, 1.17	Hγ 1.18; Hδ 0.41, 0.26
Leu[80]	8.08	4.23	1.28, 0.87	Hγ 1.37; Hδ 0.52, 0.46
Phe[85]	8.05	4.18	3.19, 2.73	Hδ 7.17; Hε 7.35, Hζ 7.02
His[86a]	8.10			Hδ$_2$ 7.38; Hε$_1$ 6.09
His[90a]	7.95			Hδ$_2$ 7.83; Hε$_1$ 6.95
Leu[91]	7.06	4.02	0.87, 0.19	Hγ 0.42; Hδ 0.60, 0.53
Ala[95]	8.66	5.03	1.30	
Phe[99]	9.56	5.53	3.05, 2.84	Hδ 7.18; Hε 7.07, Hζ 6.89
Tyr[102]	9.10	5.56	3.02, 2.74	Hδ 6.95; Hε 6.55
Val[104]	7.49	3.79	1.68	Hγ 0.45; −0.09
Trp[106]	6.95	4.62	3.49, 3.38	Hδ$_1$ 7.23; Hε$_3$ 7.70; Hζ$_3$ 7.21; Hη$_2$ 7.21; Hζ$_2$ 7.48

[a] Tentative assignment.

ization indeed involves the proteinase-interacting region of the molecule, and, moreover, that papain is not capable of breaking the cystatin C dimers.

DISCUSSION

Despite vast amounts of experimental data obtained in the last 2 decades about protein folding, rather little is known about specific intermolecular interactions, competing with the intramolecular interactions in the process of folding. To a large extent, that situation follows methodological approaches of using techniques that characterize proteins in a global way (such as CD, fluorescence, light scattering, hydrodynamic, and enzymatic methods). Results of the present work show that with the addition of the high resolution NMR spectroscopy small proteins with a limited degree of self-association or aggregation can be characterized in great detail, especially in quite common cases where small oligomers are formed. NMR spectroscopy in such a case can provide a picture of simultaneous folding and self-association or aggregation with a resolution at the amino acid level. Two other methods, agarose gel electrophoresis and size exclusion chromatography, were selected to complement the NMR studies by giving direct global information about the presence and size of the self-associated forms. Standard spectroscopic methods used to follow protein folding such as fluorescence and CD were avoided; as in the case of cystatin C, they give a complex blend of information about both self-association and folding down to micromolar concentrations.[4]

In the present work, cystatin C properties were examined under conditions close to those leading to unfolding of the protein. Three different ways of denaturing the protein were used, *i.e.* by decreasing pH, adding a denaturing reagent (guanidine hydrochloride), and raising the temperature. In all three approaches, agarose gel electrophoresis demonstrated broad pretransitional regions, in which self-association in the form of a dimerization takes place. Although some precipitation was observed, especially at high temperatures and concentrations above 1 mg/ml, no stable intermediate-sized self-associated forms were detected by either electrophoresis or SEC experiments. Pretransitional regions were examined in a more detailed way using NMR spectroscopy. In the experiments with varying pH and GdnHCl concentrations, NMR spectra directly revealed under which conditions unfolding of the protein was

occurring. All NMR results clearly indicated that the pretransitional region observed for cystatin C is completely different from the molten globule type of an intermediate state, which is quite common in protein folding paths and is often associated with an increased aggregation (Dolgikh *et al.*, 1981; Kuwajima, 1989; Ptitsyn *et al.*, 1990; Ghélis and Yon, 1982). NMR spectra of proteins in a molten globule state show features characteristic of a loss of the tertiary structure, as was well characterized for α-lactoglobulin (Dolgikh *et al.*, 1985; Baum *et al.*, 1989). Such spectra are more similar to those of the unfolded than to the native protein. On the contrary, NMR spectra of the dimeric cystatin C are very similar to those of the native monomeric form, with the majority of the tertiary structure clearly preserved. Although at the current stage of the spectral analysis there is a possibility of changes in other parts of the molecule, the majority of observed shifts in proton signals clusters around the proteinase binding site. Therefore, one can postulate that in approaching the unfolding point, the structure of the protein becomes more loose, which leads to conformational or hydration changes in the loops forming the proteinase binding site. The most evident differences in one-dimensional NMR spectra between the monomeric and dimeric form of cystatin C are in the second hairpin loop (for residues Tyr[102], Val[104], and Trp[106]), which according to computer docking of chicken cystatin with papain (Bode *et al.*, 1988) should be within van der Waals contact with the proteinase. Either conformational changes in the active site directly, or dimer formation via the binding site loops, can explain the observed loss of the activity of cystatin C as an inhibitor of cysteine proteinases upon dimerization. The K_i value for cystatin C interacting with papain is in the picomolar range (Abrahamson *et al.*, 1986), and the equilibrium constant for the self-association of cystatin C is in the micromolar range according to our present results. Still, dimerized cystatin C was unable to inhibit papain even after prolonged incubations, which must be explained by the cystatin being trapped in a dimeric form that is separated by a high energy barrier from the monomeric state (details about the barriers will be presented elsewhere).

The aggregation and precipitation during refolding of proteins produced as inclusion bodies constitutes one of the major obstacles in efficient production of recombinant proteins in

bacterial systems (Cleland, 1991; Mitraki and King, 1989). These problems were studied extensively (Ghélis and Yon, 1982; Mitraki *et al.*, 1987; Brems, 1988; Lehrman *et al.*, 1991; Cleland and Wang, 1990), with the general conclusions that folding intermediates with a partially exposed hydrophobic core are responsible for the aggregation and precipitation (Wetzel, 1992; Mitraki and King, 1989). The dimerization of cystatin C (leading to precipitation at higher concentrations) resembles other systems by exhibiting a characteristic trough (Ghélis and Yon, 1982) under conditions directly preceding unfolding. However, there is a distinct difference, as both the tertiary and secondary structure of cystatin C are conserved under such conditions. As mentioned above, the unusual large hydrophobic patch on the protein surface shows large changes in the NMR signals and, likely, participates directly in the self-association event.

It is clear from the results of the present study that cystatin C can undergo dimerization under conditions when the protein seemingly has a "normal" conformation, quite far from those leading to unfolding. It is possible that other proteins which self-associate or display problems during refolding may follow this case directly. Potentially pretransition-related changes were reported recently for the reverse transcriptase from HIV (Wright *et al.*, 1994), and, in the past, many other proteins were reported having predenaturational changes (Privalov, 1979). Similar mechanisms of interaction could play a role in folding of multidomain or oligomeric proteins. Further studies are needed to characterize such a way of self-association, and cystatin C seems to be an ideal model case for such work.

A mutated variant of cystatin C, L68Q-cystatin C, has been identified as cause for the genetic disease that is known as hereditary cystatin C amyloid angiopathy or cerebral hemorrhage with amyloidosis, Icelandic type. The Leu^{68} → Gln substitution in cystatin C results in massive systemic deposits of the protein, in particular in brain blood vessels, leading to hemorrhages and early death (reviewed by Jensson *et al.* (1987)). Recently, L68Q-cystatin C has been produced by *E. coli* expression (Abrahamson and Grubb, 1994), and it was shown that the mutated variant precipitates rapidly at human body temperatures, with transitory formation of dimers. Although more work on folding of the L68Q variant of cystatin C is needed, most likely formation of its dimers is similar to that observed in the pretransitional region for the wild-type protein. However, dimers of L68Q-cystatin C are formed at temperatures nearly 30 degrees lower than needed for the wild-type cystatin (Abrahamson and Grubb, 1994). Therefore, one can postulate that the Leu^{68} → Gln substitution lowers the transition temperature for the unfolding. That would be expected, as it is common that a substitution of an amino acid in the hydrophobic core of a protein decreases its stability (Pacula and Sauer, 1989; Alber, 1989). Results presented in this work suggest that cystatin C provides another system where decreased stability of a mutant protein correlates with its amyloidogenic nature. Increased propensity for the self-association in the broad range before unfolding may have direct relevance for the formation of the amyloid. The wild-type protein should be quite stable under physiological conditions; however, the L68Q variant should be in the self-association prone, pretransitional region. It remains to be determined if partial unfolding of the protein is necessary for the nucleation step and if dimers are indeed on the path to the fibril formation. Since we have found similarities between the behavior of L68Q and wild-type cystatin C under partially denaturing conditions, the next relevant question to address must be a structural comparison between the dimers and exploration of the role of these dimers as potential intermediates in the amyloid formation.

It also remains to be clarified if dimer formation has any physiological significance for wild-type cystatin C, especially in pathological states. The most likely possibility for a physiological dimerization would be in acidic compartments in cells, involved in the endo- and exocytosis. For example, pH in lysosomes is in the 4.6–5.0 region, which corresponds to the range where cystatin C dimerizes to some extent according to our *in vitro* results. Even lower pH values, corresponding to those resulting in extensive dimerization and parallel inactivation of cystatin C *in vitro* have been observed beneath adherent macrophages and osteoclasts (Silver *et al.*, 1988). It is important to keep in mind that our *in vitro* experiments demonstrate that cystatin C is capable of "remembering its history," as the dimeric inactive form is very easily trapped for amounts of time practically infinite on the physiological time scale. In this context, it is intriguing that inactive cystatin C dimers quite likely similar to those described in this paper are present in lysates of human neuroendocrine cells in which cystatin C is stored in secretory granules together with neuropeptides.[5] Also, cathepsins B and L have been reported to be responsible for the extracellular process of bone resorption (Delaissé *et al.*, 1984; Kakegawa *et al.*, 1993), even though cystatin C is ubiquitous extracellularly and a potent inhibitor of cathepsin B and L activity (Abrahamson *et al.*, 1986, 1990). This apparent conflict could be explained by a dimerization-caused inactivation of cystatin C in the sealed off and acidified compartment under the bone-degrading osteoclasts.

REFERENCES

Abrahamson, M., and Grubb, A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 1416–1420
Abrahamson, M., Barrett, A. J., Salvesen, G., and Grubb, A. (1986) *J. Biol. Chem.* **261,** 11282–11289
Abrahamson, M., Dalbøge, H., Olafsson, I., Carlsen S., and Grubb, A. (1988) *FEBS Lett.* **236,** 14–18
Abrahamson, M., Olafsson, I., Palsdottir, A., Ulvsbäck, M., Lundwall, Å., Jensson, O., and Grubb, A. (1990) *Biochem. J.* **268,** 287–294
Alber, T. (1989) *Annu. Rev. Biochem.* **58,** 765–798
Baum, J., Dobson, C. M., Evans, P. A., and Hanley, C. (1989) *Biochemistry* **28,** 7–13
Bax, A., and Davis, G. (1985) *J. Magn. Res.* **65,** 355–360
Berti, P. J. (1993) *Cysteine Proteases: Interaction with Cystatin C and Sequence to Structure-Function Relationships.* Ph.D. thesis, McGill University, Montreal, Canada
Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988) *EMBO J.* **7,** 2593–2599
Brader, M. L., and Dunn, M. F. (1991) *Trends Biochem. Sci.* **16,** 341–345
Brange, J., Hansen, J. F., Havelund, S., and Melberg, S. G. (1987) in *Advanced Models for the Therapy of Insulin-dependent Diabetes* (Brunetti, P., and Waldhäusl, W. K., eds) pp. 85–90, Raven Press, Rome
Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Res.* **53,** 521–528
Brems, D. N. (1988) *Biochemistry* **27,** 4541–4546
Cleland, J. L. (1991) *Am. Chem. Soc. Symp.* **470,** 1–20
Cleland, J. L., and Wang, D. I. (1990) *Biochemistry* **29,** 11072–11078
Cleland, J. L., Powell, M. F., and Shire, S. J. (1993) *Crit. Rev. Ther. Drug Carrier Systems* **10,** 307–377
Colon, W., and Kelly, J. W. (1992) *Biochemistry* **31,** 8654–8660
Dalbøge, H., Bech Jensen, E., Tøttrup, H., Grubb, A., Abrahamson, M., Olafsson, I., and Carlsen, S. (1989) *Gene (Amst.)* **79,** 325–332
Delaissé, J.-M., Eeckhout, Y., and Vaes, G. (1984) *Biochem. Biophys. Res. Commun.* **125,** 441–447
De Young, L. R., Fink, A. L., and Dill, K. A. (1993) *Acc. Chem. Res.* **26,** 614–620
Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S. Yu., and Ptitsyn, O. B. (1981) *FEBS Lett.* **136,** 311–315
Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bychkova, V. E., Gilmanshin, P. I., Lebedev, Yu. O., Semisotnov, G. V., Tiktopulo, E. I., and Ptitsyn, O. B. (1985) *Eur. Biophys. J.* **13,** 109–121
Engh, R. A., Dieckmann, T., Bode, W., Auerswald, E. A., Turk, V., Huber, R., and Oschkinat, H. (1993) *J. Mol. Biol.* **234,** 1060–1069
Frangione, B. (1991) in *Amyloid and Amyloidosis 1990* (Natvig, J. B., Førre, Ø., Husby, G., Husebekk, A., Skogen, B., Sletten, K., and Westermark, P., eds) pp. 351–356, Kluwer Academic Publishers, Dordrecht
Ghélis, C., and Yon, J. (1982) *Protein Folding*, pp. 289–295, Academic Press, New York
Ghiso, J., Jensson, O., and Frangione, B. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 2974–2978
Glatz, C. E. (1992) in *Stability of Protein Pharmaceuticals* (Ahern, T. J., and Manning, M. C., eds) Part A, pp. 135–166, Plenum Press, New York
Glenner, G. G., Mehlhaff, P. M., and Kawano, H. (1991) in *Amyloid and Amyloidosis 1990* (Natvig, J. B., Førre, Ø., Husby, G., Husebekk, A., Skogen, B.,

---

[5] C. Vogel and M. Abrahamson, unpublished observations.

Sletten, K., and Westermark, P., eds) pp. 707–712, Kluwer Academic Publishers, Dordrecht

Hall, A., Håkansson, K., Mason, R. W., Grubb, A., and Abrahamson, M. (1995) *J. Biol. Chem.* **270,** 5115–5121

Honig, B., and Yang, A.-S. (1995) *Adv. Prot. Chem.* **46,** 27–58

Hurle, M. R., Helms, L. R., Li, L., Chan, W. N., and Wetzel, R. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 5446–5450

Jensson, O., Gudmundsson, G., Arnason, A., Blöndal, H., Petursdottir, I., Thorsteinsson, L., Grubb, A., Löfberg, H., Cohen, D., and Frangione, B. (1987) *Acta Neurol. Scand.* **76,** 102–114

Jeppsson, J.-O., Laurell, C.-B., and Franzén, B. (1979) *Clin. Chem.* **25,** 629–638

Jovin, T. M. (1973) *Ann. N. Y. Acad. Sci.* **209,** 477–496

Kakegawa, H., Nikawa, T., Tagami, K., Kamioka, H., Sumitani, K., Kawata, T., Drobnic-Kosorok, M., Lenarčič, B., Turk, V., and Katunuma, N. (1993) *FEBS Lett.* **321,** 247–250

Kuwajima, K. (1989) *Proteins: Struct. Funct. Genet.* **6,** 87–103

Lehrman, S. R., Tuls, J. L., Havel, H. A., Haskell, R. J., Putnam, S. D., and Tomich, C.-S. (1991) *Biochemistry* **30,** 5777–5784

McCutchen, S. L., Colon, W., and Kelly, J. W. (1993) *Biochemistry* **32,** 12119–12127

Mitraki, A., and King, J. (1989) *Bio/Technology* **7,** 690–697

Mitraki, A., Betton, J.-M., Desmadril, M., and Yon, J. M. (1987) *Eur. J. Biochem.* **163,** 29–34

Pacula, A. A., and Sauer, R. T. (1989) *Annu. Rev. Genet.* **23,** 289–310

Palsdottir, A., Abrahamson, M., Thorsteinsson, L., Arnason, A., Olafsson, I., Grubb, A., and Jensson, O. (1988) *Lancet* **2,** 603–604

Privalov, P. L. (1979) *Adv. Protein Chem.* **33,** 167–241

Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Žerovnik, E., and Razgulyaev, O. I. (1990) *FEBS Lett.* **262,** 20–24

Silver, I. A., Murrills, R. J., and Etherington, D. J. (1988) *Exp. Cell Res.* **174,** 266–276

Sipe, J. D. (1992) *Annu. Rev. Biochem.* **61,** 947–975

Stone, M. J. (1990) *Blood* **75,** 531–545

Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarčič, B., and Turk, V. (1990) *EMBO J.* **9,** 1939–1947

Wetzel, R. (1992) in *Stability of Protein Pharmaceuticals* (Ahern, T. J., and Manning, M. C., eds) Part B, pp. 43–88, Plenum Press, New York

Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol.* **239,** 363–392

Wright, L. L., Wilson, J. E., Martin, J. L., Haire, S. E., Ray, P. H., and Painter, G. R. (1994) *Biophys. J.* **66,** A181

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp. 13–31, John Wiley & Sons, New York

Folding-related Dimerization of Human Cystatin C

Irena Ekiel and Magnus Abrahamson

Access the most updated version of this article at http://www.jbc.org/content/271/3/1314

Alerts:

- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 44 references, 8 of which can be accessed free at
http://www.jbc.org/content/271/3/1314.full.html#ref-list-1