3D Domain-Swapped Human Cystatin C With Amyloidlike Intermolecular β-Sheets

Robert Janowski,[1,2] Maciej Kozak,[1,3] Magnus Abrahamson,[4] Anders Grubb,[4] and Mariusz Jaskolski[1,2]*

[1]*Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland*
[2]*Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland*
[3]*Department of Macromolecular Physics, A. Mickiewicz University, Poznan, Poland*
[4]*Department of Clinical Chemistry, Lund University, Sweden*

ABSTRACT **Oligomerization of human cystatin C (HCC) leads to amyloid deposits in brain arteries, and this process is greatly accelerated with a naturally occurring L68Q variant. The crystal structures of N-truncated and full-length HCC (cubic form) showed dimer formation via three-dimensional (3D) domain swapping, and this observation has led to the suggestion that an analogous domain-swapping mechanism, but propagated in an open-ended fashion, could be the basis of HCC fibril formation. Here we report that full-length HCC, when crystallized in a new, tetragonal form, dimerizes by swapping the same secondary structure elements but with a very different overall structure generated by the flexibility of the hinge linking the moveable elements. The β-strands of the β-cores of the two folding units of the present dimer are roughly parallel, while they formed an angle of about 100° in the previous two structures. The dimers pack around a crystallographic dyad by extending their molecular β-sheets in an intermolecular context. At the other edge of the molecular β-sheet, side-chain–side-chain hydrogen bonds propagate the β-structure in the same direction. In consequence, a supramolecular crystal structure is generated, with all the β-strands of the domain-swapped dimers being perpendicular to one crystallographic direction. This observation is relevant to amyloid aggregation of HCC, as X-ray diffraction studies of amyloid fibrils show them to have ordered, repeating structure, consistent with the so-called cross-β structure, in which extended polypeptide chains are perpendicular to the fiber axis and form infinite β-sheets that are parallel to this axis. Proteins 2005; 61:570–578.** © **2005 Wiley-Liss, Inc.**

Key words: human cystatin C; 3D domain swapping; cross-β structure; amyloid fibril; cysteine protease inhibitor

INTRODUCTION

Human cystatin C (HCC) is involved in the pathophysiology of two types of amyloid disorders, namely, in hereditary cystatin C amyloid angiopathy (HCCAA), in which an L68Q mutant is deposited as amyloid, causing brain hemorrhage in early adulthood,[1] and in disorders involving deposition of amyloid β fibrils with wild-type (wt) cystatin C as a coprecipitant.[2] The single polypeptide chain of HCC consists of 120 amino acid residues and includes four cysteine residues in the C-terminal part, paired into disulfide bonds 73–83 and 97–117.[3,4] Physiologically, monomeric HCC functions as a potent inhibitor of cysteine proteases of the papain and legumain families.[5] The inhibitory epitope for proteases of the papain family consists of three elements aligned at one edge of the molecule,[6] namely, the N-terminal peptide and two hairpin loops, L1 and L2. These two loops are elements of a large and curved five-stranded antiparallel β-sheet that, together with a long perpendicular α-helix forms a β-grip motif. The connectivity within this topology is (N)-β1-α-β2-L1-β3-AS-β4-L2-β5-(C), where AS is a broad "appending structure" positioned at the opposite ("back-side") end of the β-sheet relative to the N/L1/L2 edge. The AS structure harbors the epitope involved in the inhibition of cysteine proteases of the legumain family.[7]

Whereas L68Q cystatin C undergoes dimerization and oligomerization at the temperature of the human body and is found in dimeric form in vivo, wt cystatin C shows higher stability,[8,9] but it can also be dimerized at elevated temperature, low pH, or at conditions of mild chemical denaturation.[10,11] The structure of wt HCC determined for cubic crystals grown at acidic pH (4.8) showed dimer formation via the mechanism of three-dimensional (3D) domain swapping.[12] In 3D domain swapping, which is a mechanism of protein oligomerization,[13,14] an exchangeable structural element of one protein molecule is anchored in a mutual or circular fashion in another protein molecule in the same way as in the folding of a single protein chain. In HCC, the swapped element consists of the α-helix and the two flanking strands, β1 and β2. Residue 68 is located at the C-terminal end of strand β3, and its side-chain is anchored in a hydrophobic pocket

Grant sponsor: State Committee for Scientific Research; Grant number: 4 T09A 039 25 (to M. Jaskolski). Grant sponsor: Foundation for Polish Science. Grant sponsor: Swedish Science Council; Grant numbers: 05196 and 09915. Grant sponsor: Medical Faculty of Lund University. Grant sponsors: Crafoord, A. Påhlsson, A. Österlund, and G. and J. Kock Foundations.

*Correspondence to: Mariusz Jaskolski, Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Grunwladzka 6, 60-780 Poznan, Poland. E-mail: mariuszj@amu.edu.pl

Received 7 February 2005; Accepted 6 May 2005

Published online 16 September 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20633

formed by residues at the C-terminal end of the α-helix. The pocket is sufficient to accommodate a residue as large as leucine, but not glutamine, which explains the lower stability of the L68Q mutant. In addition, after partial unfolding of the protein, the exposed hydrophilic Q68 patch of the mutant is less unstable in contact with solvent, thus further promoting a monomer→partially unfolded state→aggregate transformation.

Since the protein extracted from cystatin C amyloid is shortened at the N-terminus,[15] it was of interest to see if N-truncated HCC (THCC) is also capable of 3D domain swapping. The crystal structure of THCC showed virtually identical domain-swapped dimers with only slight variations of the open interface, which consists of a dimer-specific (i.e., not present in the monomer) β-sheet structure formed by the open L1 loops.[16]

The crystallographic data, combined with the observation of dimeric HCC in the blood plasma of HCCAA patients,[9] have led to the suggestion that an analogous 3D domain-swapping process, but propagated in an open-ended fashion, could be the basis of the formation of cystatin C amyloid fibrils.[12,17] This notion has been reinforced by the experiments reported by Nilsson et al.,[11] who showed that strategic placement of new disulfide bridges in the monomeric fold of cystatin C resulted in total inhibition of dimerization and greatly reduced the amyloid formation process. The engineered S—S bonds were introduced either between strands β2 and β3 or between the α-helix and strand β5 (i.e., always in such a way as to prevent domain swapping) by linking elements of the swapped domain (β2, α) with the fixed core. An independent support for the hypothesis implicating 3D domain swapping in amyloidogenesis came from the crystallographic observation of domain-swapped dimers of the human prion protein,[18] and from the elegant follow-up experiments demonstrating in vitro fibrillar aggregation and infectivity for a recombinant prion protein.[19]

X-ray diffraction patterns of amyloid fibrils show them to have ordered, repeating structure, consistent with the so-called cross-β structure,[20,21] in which extended polypeptide chains in β-conformation are perpendicular to the fiber axis, and form β-sheets that are parallel to the fiber axis. However, amyloid fiber diffraction experiments provide structural information only at a very limited level, not sufficient for example to confirm, or disprove, the presence of 3D domain swapping. It is, therefore, of interest to see if aggregation of amyloidogenic proteins could be detected beyond the closed-ended dimeric form, for example, by X-ray crystallography.

Here we report the crystal structure of a new, tetragonal, polymorphic form of full-length HCC and show that the molecules again have dimerized by swapping the β1-α-β2 fragment. However, the overall shape of the dimer is very different from those observed earlier, demonstrating that the hinge region, which in dimeric HCC forms the open interface between the two polypeptide chains, has a high degree of plasticity, allowing structural adaptation to different molecular environments. More importantly, the new dimers enter into strong intermolecular interactions involving the free edges of the molecular β-sheet. In particular, parallel β5–β5 interactions between two domain-swapped dimers lead to the creation of a tetrameric unit. The tetramers interact via side-chain–side-chain hydrogen bonds at their β2 strands leading to infinite aggregation in which all the β-chains in the crystal are roughly perpendicular to the **c** direction. This situation, where interactions of β-chains create an infinite network encompassing all protein molecules in the crystal, is reminiscent of the crystal packing reported for a llama VHH domain (antigen-binding fragment of the heavy-chain antibody), in which the crystallographic building block was also a domain-swapped dimer.[22] The uniqueness of the present structure, however, lies in the perpendicular organization of all the β-chains relative to a common direction.

TABLE I. Data Processing and Structure Refinement Statistics

Space group	$P4_12_12$
Cell parameters (Å)	$a = 91.5, c = 144.5$
Temperature (K)	120
Resolution (last shell) (Å)	40.0–3.03 (3.08–3.03)[a]
Measured reflections	161,313
Unique reflections	12,273
Completeness (%)	97.8 (97.7)
$<I/\sigma(I)>$	14.0 (3.2)
R_{int}[b]	0.123 (0.409)
Refinement program	CNS
Resolution range (Å)	15.0–3.03
Number of reflection used	11,902
R[c]	0.217
R_{free}[d]	0.262
Number of protein atoms	1776
Water molecules	22
RMSDs from ideality:	
Bond lengths (Å)	0.017
Bond angles (°)	2.0
Ramachandran plot statistics (%)	
Most favored regions	88.9
Allowed regions	11.1

[a]Values in parentheses correspond to the last resolution shell.
[b]$R_{int} = \Sigma_h\Sigma_j |I_{hj} - <I_h>| / \Sigma_h\Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h.
[c]$R = \Sigma_h||F_o| - |F_c||/\Sigma_h|F_o|$ for all reflections, where F_o and F_c are observed and calculated structure factors, respectively.
[d]R_{free} was calculated against 10% of all reflections randomly excluded from the refinement.

MATERIALS AND METHODS

Crystallization and Data Collection

Crystallization of full-length HCC has been reported by Kozak et al.[23] The present study is based on the tetragonal form crystallized from solutions of monomeric protein prepared by size-exclusion chromatography in the final purification step. Low-temperature X-ray diffraction data were measured to 3 Å resolution using the BW7A EMBL beamline of the DESY synchrotron. Cryoprotection was provided by the mother liquor composition, which included 50–60% of 2-methyl-2,4-pentanediol (MPD). The data were integrated and scaled in the HKL package.[24] The data-processing statistics are shown in Table I.

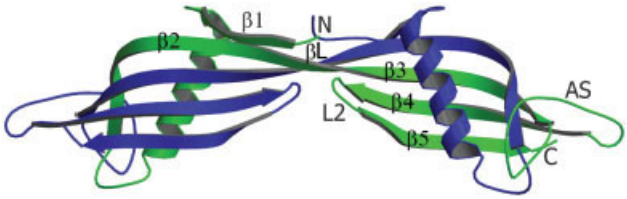

Fig. 1. Human cystatin C dimer with swapped domains present in the asymmetric unit of the tetragonal crystal form of the protein. The two polypeptide chains are shown in different colors. The β-strands forming the antiparallel β-sheet in each domain are numbered from β1 (top strand) to β5 (bottom strand). They form a grip around the central α-helix, which runs across the face of the β-sheet. Note that strands β2 traverse from one domain to the other (where they continue as β3) through the βL segment (open interface), which is created by opening of loop L1 present in the monomeric molecule. The second hairpin loop of the monomer, L2, is intact and seen close to the dimer interface. Also the N-terminal chains are directed toward the dimer interface. Some residues are missing from the amino-ends of the polypeptide chains due to disorder, although the crystallized protein contained the complete sequence. The unstructured loops on the periphery of the molecule are termed "appending structure" (AS), and are not affected by the dimerization process. Figure prepared in MOLMOL.[30]

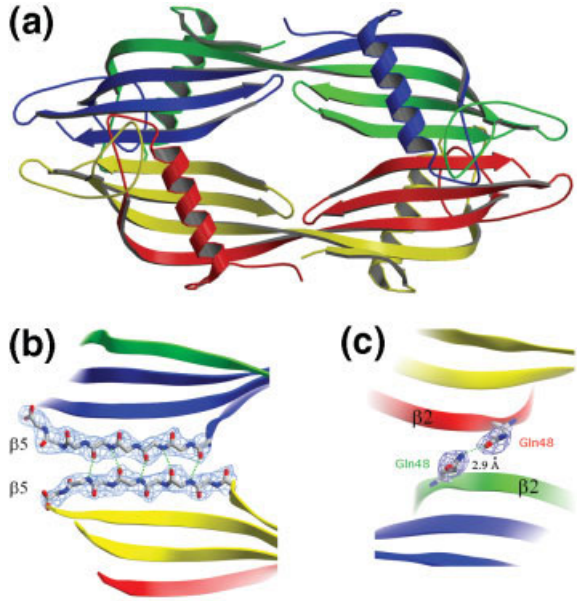

Fig. 2. Interactions of the 3D domain-swapped dimers of tetragonal HCC at the edges of the molecular β-sheets. An HCC tetramer, formed from two dimers (AB, blue and green, and a symmetry-related dimer A′B′, red and yellow) via two systems of parallel β-interactions between the β5-strands (**a**). Each of the β5–β5 motifs, shown in a $2F_o$-F_c map (contour level 1.1σ), includes five main-chain–main-chain hydrogen bonds (**b**). At the other edges of the β-sheets, neighboring dimers interact via side-chain–side-chain hydrogen bonding between their β2-chains, shown in a $2F_o$-F_c map contoured at the 1.1σ level (**c**). Figure prepared in MOLMOL[30] and in CueMol.[36]

Structure Determination and Refinement

The crystal structure was solved by molecular replacement using MOLREP.[25] As the molecular probe, one-half (i.e., the folding unit) of the crystallographic HCC dimer (cubic form) was used[12] [Protein Data Bank (PDB) code: 1G96]. The molecular replacement solution identified two copies of the search model, packed together to form a dimer consisting of intertwined polypeptide chains of two HCC molecules. The structure was refined using the maximum-likelihood option of the Crystallography & NMR System (CNS) program.[26] After rebuilding in O,[27] the final model, refined with all 11,902 reflections between 15 and 3.03 Å, is characterized by an R factor of 0.217 and R_{free} of 0.262. A few residues of each N-terminal peptide could not be modeled because of disorder. Except for the N-termini, the polypeptide main-chain can be traced without breaks in the $2F_o$-F_c map. Also the side-chains (with a few exceptions at the molecular surface) have very clear definition. The very good quality of the electron density maps allowed the inclusion of 22 perfectly defined water molecules. The structure–refinement statistics are shown in Table I. The coordinates and structure factors of the tetragonal form of HCC have been deposited with the PDB under the accession code 1TIJ.

RESULTS AND DISCUSSION

Structure Solution and Refinement

The molecular replacement calculations produced a plausible solution in the $P4_12_12$ space group (R factor 48.7%, correlation coefficient 44.1%) revealing two copies of the search probe (i.e., two units with cystatin fold) in the asymmetric unit, located in close proximity and consistent with covalent connection. Manual modeling of the molecular replacement solution confirmed the presence of two HCC molecules, intertwined into a 3D domain-swapped dimer. Such a composition of the asymmetric unit corresponds to a Matthews volume of 5.60 $Å^3$ Da^{-1} and to a very high solvent content (77.9%).[28] Careful examination of additional molecular replacement results, obtained after inclusion of the partial model, as well as inspection of the nonprotein regions of the unit cell, confirmed beyond doubt that no other copies of HCC were present. The high solvent content of the present crystals explains their poor diffraction power and is reminiscent of the solvent volume in the cubic form of full-length HCC (71.5%).[12] In contrast, the solvent content reported for the N-truncated form of cystatin C is quite normal (51.4%).[16] It is conceivable, that it is the flexible N-terminal decapeptide, which even if present in the sequence tends to be disordered in the crystal structure, that prevents the HCC molecules from tight crystal packing.

The final model has very good geometry, characterized by an root-mean-square deviation (RMSD) from ideal bond lengths of 0.017 Å. Stereochemical analysis performed using PROCHECK[29] indicates that there are no residues with generously allowed or unfavorable backbone dihedral angles and that 88.9% of all residues are in the core region of the Ramachandran plot. The N-terminal segment of both polypeptide chains shows high mobility, and disorder and is not visible in the electron density maps (six residues in chain A and eight residues in chain B). An analogous situation was observed in the cubic crystal form of the full-length protein, where nine N-terminal residues were missing from the electron-density-derived model.[12]

TABLE II. Comparison of the Main-Chain Torsion Angles (°) in the Linker Region (Given for Each Residue in the Order: ϕ ψ ω) in the Present Structure, in Cubic HCC (PDB Code: 1G96), and in the Energy-Minimized Average NMR Model of Domain-Swapped Cystatin A (PDB Code: 1N9J)

Residue		R53	K54	Q55	I56	V57	A58	G59	V60	N61
This work	A	−107 155 176	−128 175 −178	−161 159 −179	−125 113 176	−82 129 179	−115 141 177	−129 118 177	−101 129 174	−99 134 179
	B	−111 155 178	−137 154 178	−146 132 −179	−105 116 179	−109 129 −175	−129 170 −177	−145 161 −178	−122 131 175	−104 125 178
1G96		−137 152 −178	−136 160 −179	−157 129 179	−87 131 −177	−127 153 179	−135 148 179	−158 −177 −179	−128 130 176	−95 129 175
Cystatin A 1N9J (NMR)		−128 173 −177 −128 155 177 K44	−143 128 −179 −95 162 −174 T45	−113 133 −175 −135 135 179 Q46	−109 124 175 −109 118 162 V47	−90 145 176 −85 151 172 V48	−122 148 165 −127 143 166 A49	−128 174 −172 −123 152 −172 G50	−129 115 175 −124 111 161 T51	−108 147 −170 −93 156 −175 N52

The correspondence of the residues between cystatin C and cystatin A (given in the bottom row) is based on sequence alignment[33] correlated with structural alignment.[34] For the present dimer and the cystatin A dimer, which possess only approximate symmetry, the values for both protein chains (A and B) are given.

TABLE III. Superpositions[34] of the Present HCC structure ("Superposed Entity") on the Existing Crystallographic Models of Cystatin C and on the NMR Model of Cystatin A

Superposed entity		HCC-tetragonal (this work)	HCC-cubic[12]	THCC[16] CD	THCC[16] GH	Cystatin A[35] (NMR)
Dimer	AB	—	4.73	4.80	4.72	4.29
	BA	1.85 (180.0)	—	5.00	4.64	4.24
Monomer	A	—	4.71	4.81/4.51	4.75/4.78	3.91/3.68
	B	1.80 (179.0)	4.53	4.43/4.49	4.47/4.37	3.89/3.65
Domain	$\frac{1}{2}$(AB)	—	0.91	0.96/0.81	0.79/0.87	3.04/3.98
	$\frac{1}{2}$(BA)	0.36 (170.4)	0.72	0.92/0.90	0.96/0.89	3.07/3.96

The results are given as RMSDs (Å) between the superposed Cα atoms and, in parentheses, as the angle of rotation (°), where appropriate. The superpositions were calculated for the complete dimer in two orientations (denoted AB and BA), for the individual chains (chain-on-chain), and for the two structural domains (domain-on-domain) of the dimer. For N-truncated cystatin C, THCC (for which two, CD and GH, out of four, representative domain-swapped dimers have been selected), and for cystatin A, where the domain swapped dimers are only pseudosymmetric (in contrast to cubic HCC), the calculations (where appropriate) are reported for both components of the target.

The Domain-Swapped Dimer

The two cystatin C chains are intertwined to form a domain-swapped dimer (Fig. 1). Each of the structural domains (folding units) has the canonical cystatin fold,[6,31] also found in the cystatin C structures described earlier[12,16] [i.e., follows the pattern (N)-β1-α-β2-L1-β3-AS-β4-L2-β5-(C)]. The structural element that is swapped, in a reciprocated fashion, in the act of dimerization is the same as in the other domain-swapped dimers of cystatin C[12,16] [i.e., consists of the (N)-β1-α-β2 motif from the N-terminal part of the molecule]. In consequence, the closed interface[32] is the same as in the other dimers of HCC and as in the monomeric molecule. However, the linker region (i.e., the L1 loop, which undergoes a conformational change between the monomeric and dimeric form) has a very different conformation from that found in the other HCC dimers. In particular, Table II shows that transition from the form observed in the cubic crystal to the present conformation requires a change of about 30° the backbone ϕ angle of residue I56. The result is an entirely different shape of the dimeric molecule. While the linker region in the open interface of cubic HCC[12] and of THCC[16] had a curved shape and positioned the prominent β-strands in the two folding domains at an angle of about 100°, in the present dimer it runs from one domain to the other in a straight fashion, resulting in a nearly parallel orientation of all the β-strands (Fig. 1). The two polypeptide chains composing the dimer can be superposed[34] using a rotation of 179.0°, showing a general preservation of the two-fold symmetry present in the cubic structure of HCC. However, the agreement is less perfect in terms of RMSD between the superposed Cα atoms, which is as high as 1.80 Å (Table III). This is also visible in a comparison of the backbone torsion angles in the linker area between chains A and B. Several differences are about 30° with a culmination (43°) at the ψ angle of the flexible glycine 59 residue (Table II). On the other hand, when the two folding units (domains) of the present dimer are superposed, the RMSD value is very low (0.36 Å), confirming the high degree of conservation of the folding pattern, but the angle of rotation is 170.4°.

TABLE IV. Main-Chain–Main-Chain Hydrogen Bonding Between Chains β1 and β2 of Each Monomer (A/B)

Hydrogen bond	Distance (Å)
(L9)O–N(A58)[a]	3.49/2.92
(G11)N–O(I56)[a]	—/3.17
(M14)N–O(K54)	3.00/2.91
(M14)O–N(K54)	2.77/2.89
(A16)N–O(A52)	2.65/2.75

[a]Not found in cubic HCC and in THCC.

TABLE V. Hydrogen Bonds Linking the Edges of the Molecular β-Sheets Formed Between Symmetry-Related HCC Dimers in the Present Structure

Hydrogen bond	Distance (Å)
β5–β5	
(S113)O–N(S115)	3.53
(S115)N–O(S115)	3.11
(S115)O–N(C117)	2.94
(C117)N–O(C117)	3.06
(C117)O–N(D119)	3.20
β2–β2	
(Q48)Nε2–Oε1(Q48)	2.90

The interactions of the β5-chains are between main-chains of molecules A and B′, and B and A′. The interaction of the β2 side-chains is between molecules A and B″ and between molecules B and A″′.

The overall shape of the present HCC dimer is more reminiscent of the 3D domain-swapped dimer of the related cystatin A detected by NMR spectroscopy in solution.[35] Table II illustrates that the differences between the backbone torsion angles of the present dimer and of the PDB entry 1N9J (energy-minimized average NMR structure of dimeric cystatin A) are smaller than the variations seen within each dimer and in their comparisons with the cubic HCC dimer. The fact that the same protein (HCC) is capable of dimerization via 3D domain swapping with different conformations at the open interface demonstrates the high degree of flexibility of this oligomer and its structural adaptability to the molecular environment, such as can be found in a crystal or in the amyloid fibril. Despite the different conformation of the linker region, the system of antiparallel β-interactions between the two copies of the β2-βL-β3 chain (where βL is the L1 sequence in a new conformation) is unchanged, and as in cubic HCC and THCC consists of 32 main-chain–main-chain hydrogen bonds.

The flexibility of the N-terminal segment does not allow model building starting from residue S1, even though the complete sequence is present in the crystallized polypeptide. However, in the present structure the missing segments are shorter (six and eight residues in chains A and B, respectively) than in cubic HCC (nine residues) despite a higher solvent content. The visible N-terminal stems are clearly extended beyond β1, the first strand of the β-sheet. The extensions tend toward the central area of the dimer, providing protection for the intermolecular β-sheet at the dimer interface, while leaving the N-terminal portion of strand β2 unprotected, so that it becomes in that portion the effective edge of the otherwise five-stranded molecular β-sheet (Fig. 1). The protection of β1–β2 interactions starts from residue A52 of strand β2 (Table IV). It is interesting to note that strand β1 forms hydrogen bonds also with residues corresponding to the L1 loop of monomeric cystatin C (I56, A58). Beyond this point, from residue R8, the N-terminal peptide is not attached to the protein core and thus free to move in the solvent.

The appending structure of the cystatin fold is largely unstructured and is sometimes only partially modeled. Surprisingly, in the present crystal form the two AS structures could be modeled without ambiguity in spite of the high solvent content of the crystals and limited resolution of the data. The AS segments of both monomers have very similar conformation. In agreement with the crystallographic models of cubic HCC and THCC, no helical structure is present in this segment, in variance with the model proposed for chicken cystatin.[31] The AS loops are located on the periphery of the molecule, at a maximum distance from the area where the two chains switch domains. This architecture is consistent with the observation that the capacity to inhibit human legumain, associated with the AS epitope,[7] is not affected by HCC dimerization.

Aggregation of the Dimeric Units

The domain-swapped dimers form a rich network of interactions of symmetry-related molecules. A pair of two-fold symmetry related dimers associate to form a tetramer via parallel β-sheet interactions between their β5-strands [Fig. 2(a, b)]. These interactions, which occur in the tetramer in two copies and consist of 10 main-chain–main-chain hydrogen bonds (Table V), extend the molecular β-sheets of the dimers into intermolecular context. Such extensive interactions are possible because the β5 strand is located at the edge of the molecular β-sheet and is only partially protected by the C-terminal end of the α-helix and the following loop, which extend beyond the boundary of the β-sheet (Fig. 1).

At the other edge of the β-sheet, adjacent tetramers form side-chain–side-chain interactions via their β2-strands in antiparallel orientation [Fig. 2(c)]. In this case, there is only one point of attachment between a pair of interacting tetramers. Since each tetramer has four such points of attachment, this mode of interaction propagates the lattice interactions throughout the entire crystal (Fig. 3). There is only one direct hydrogen-bond link per each β2–β2 interaction, formed by the side-chains of crystallographic copies of residue Q48 (Table V), but the symmetry-related molecules forming the β2–β2 motif have additional hydrogen-bond contacts, in particular involving residues in the β1-α loop and in the N-terminal half of the α-helix. The dimer–dimer interactions in the β2 area are weaker than the double-lock association via the β5-strands. This should be viewed as a possible consequence of packing requirements and not necessarily as an indication of an intrinsically lower capacity of the β2-strand to support intermolecular β-sheet interactions. For instance, in the structure of THCC, the dimers were found to aggregate by parallel

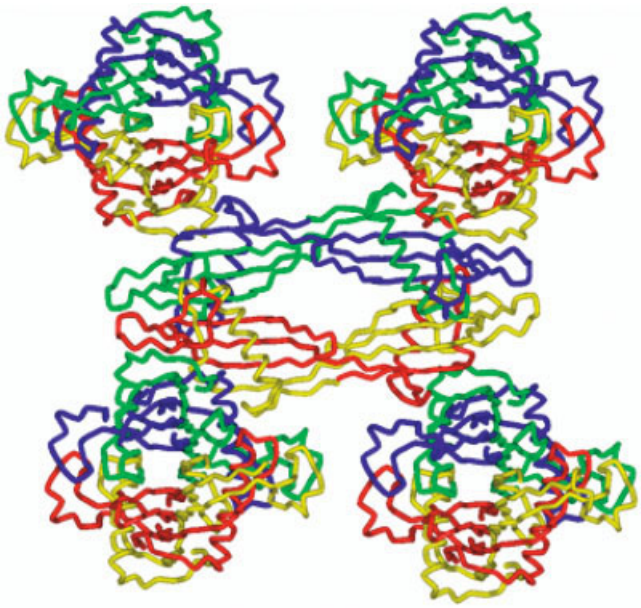

Fig. 3. Packing of the domain-swapped HCC dimers in the tetragonal crystal form. In the center, a tetramer formed from two dimers (blue/green and red/yellow) by β5–β5 interactions is shown, as in Figure 2(a). The tetramers in the upper and lower layers are connected to the central unit via β2–β2 interactions, as in Figure 2(c). This system of interactions extends in three dimentions and encompasses the entire crystal. Figure prepared in ViewerLite.[37]

β-sheet interactions between their β2 main-chains. The THCC case is very intriguing because the interacting β2-strands, forced to diverge by packing requirements, recruit a water molecule to seal a breaking β–β link.[16]

The problem of intermolecular β2–β2 interactions raises the question of intermolecular accessibility of the "free edge" of the β2-chain. Since β2 is not the first strand of the molecular β-sheet, it is naturally protected by the β1-strand and the unstructured N-terminal peptide. However, since the β1-strand is rather short and associated only with the C-terminal part of β2 (Table IV), the N-terminal portion of β2 should be still available for intermolecular interactions. In N-truncated cystatin C, where the N-terminal decapeptide is physically missing, the accessibility of strand β2 may be even better, in agreement with the solvent content and packing interactions in the THCC crystal structure.[16]

As a consequence of the symmetry-determined parallel orientation of strands β5 and antiparallel orientation of strands β2, and coupled with the generally parallel disposition of the β-structures of the domain-swapped dimer, the entire population of β-chains within the tetragonal crystal has a preferred, common orientation. In this alignment, the β-chains are roughly perpendicular to the crystallographic **c** direction and are organized into β-structures of very strong association (within the domain-swapped assemblies), of strong association (within the parallel β5–β5 structures), and of weak association (within the β2–β2 structures). Regardless of the strength of these interactions, it is remarkable that the entire crystal becomes one big agglomerate formed from 3D domain-swapped units, which aggregate further using virtually only the interactions described above (Fig. 3). Such a molecular organization is very attractive from the point of view of the molecular architecture of amyloid fibrils, which from fiber diffraction studies[38] and other evidence (e.g., cryoelectron microscopy[39]) are believed to possess the so-called cross-β structure,[20,21] in which an infinite progression of β-chains is arranged in a perpendicular fashion relative to the fiber axis. A somewhat similar situation, with β-chain interactions connecting domain-swapped protein aggregates into an infinite network encompassing all the molecules in the crystal, has been reported for a llama VHH domain.[22] The uniqueness of the present structure, however, lies in the systematically perpendicular organization of all the β-chains relative to one common direction.

Comparison With Other HCC Structures

It is instructive to compare all the crystallographic models of domain-swapped HCC. Figure 4 illustrates that the dimers are not only capable of adopting slightly different conformations at the βL–βL hinge, visible in a comparison of the cubic HCC and THCC structures,[16] but that the conformational change can be very dramatic, leading to an entirely new shape of the molecule. As mentioned above, a rough estimate of this change is obtained by comparing the mutual orientation of the β-strands in the two domains of the dimer, which are approximately perpendicular in cubic HCC and THCC, whereas in the present structure, they have parallel orientation. A more precise quantitative measure of the swing angle of one domain with respect to the other is obtained by first superposing one domain of the present AB dimer on one domain of the cubic HCC dimer, and then calculating the rotation angle which would bring into superposition the remaining domains. This value, calculated in ALIGN,[34] is equal to 69.1° and illustrates the large swinging motion necessary for interconvertion between the two conformations of the domain-swapped HCC dimer.

The monomerlike domains in all the forms of cystatin C are, however, quite similar, yielding in pairwise Cα superpositions RMSD values that do not exceed 1 Å (Table III) and are only somewhat higher than those reported for the rather similar cubic HCC and THCC dimers.[16] The similarity of the domains is not reflected in the comparisons of the individual polypeptide chains, where the RMSDs are as high as 4.8 Å (Table III). This is obviously the effect of the different conformation of the βL linker (Table II), which results in incompatible dispositions of the swapping domain with respect to the "stationary" core. As expected, the values obtained in comparisons involving the cubic HCC dimer are quite similar to those using the THCC targets (Table III).

As mentioned earlier, the present dimer bears closer similarity to the cystatin A dimer in solution than to the cubic HCC dimer (Table III). However, the resemblance is still limited. Considering the overall shape of the cystatin A dimer, it appears to be even more open through the swinging motion of one domain relative to the other than

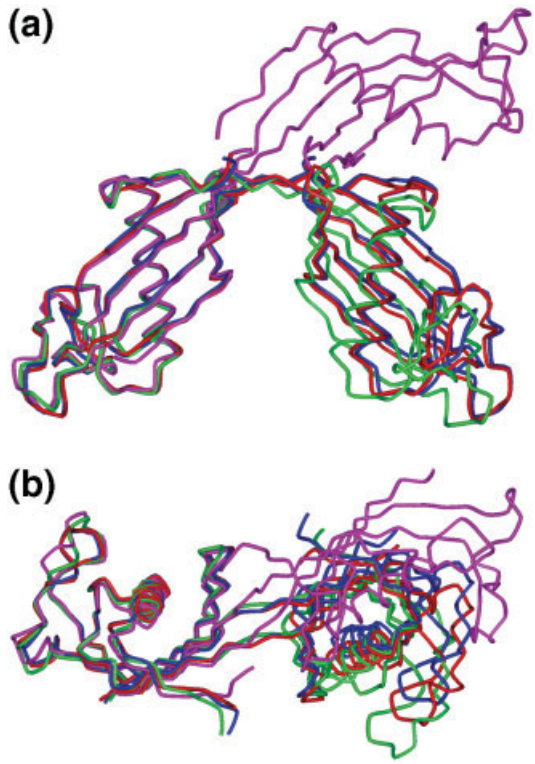

Fig. 4. Superposition of the dimeric HCC structures emphasizing the plasticity of the molecules originating from the flexibility of the hinge element at the open interface. The present full-length HCC dimer is shown in magenta; the crystallographic dimer of the cubic crystal form of HCC is shown in blue, and the dimers CD and EF of the N-truncated variant (THCC) are shown in red and green, respectively. The superpositions, shown in two perpendicular orientations, (**a**) and (**b**), were calculated in ALIGN[34] using the Cα atoms from only one-half (lefthand side) of the molecules. The differences in the dimer geometry are therefore emphasized in the righthand side domain. Figure prepared in ViewerLite.[37]

the present tetragonal dimer of HCC. The individual domains of cystatin A show a rather poor similarity to the cystatin C fold (Cα superpositions with RMSDs of 3–4 Å), not unexpected considering the different subclassification of these inhibitors.[40]

Implications for Amyloid Aggregation

Although there is growing circumstantial evidence and increasing acceptance for the hypothesis implicating 3D domain swapping in amyloidogenesis,[19,41,42] direct experimental proof is still missing. However, even if one agrees that 3D domain swapping might be indeed involved in the formation of amyloid fibrils, several fundamentally different scenarios are still possible. In one scenario, the aggregating protein could be capable of swapping two domains, one at the N- and one at the C-terminus. Such a possibility has been found for ribonuclease (RNase) A,[43,44] but in general it does not appear to be a common situation. In this scenario, the growing polymer would not have "sticky ends," or incomplete domains necessitating complementation by the next domain-swapping event; the ends would be protected and inert from the point of view of aggregation until an unfolding event exposed the domains for swapping. Therefore, kinetically and thermodynamically, the growth phase would not be very different from the nucleation phase.

In another scenario, the domain that undergoes closed-ended, mutual swapping (as in the numerous crystal structures of 3D domain-swapped oligomers) could, in principle, undergo infinite open-ended swapping, leading to linear polymerization. In this scenario, the initial and late stages of aggregation would be markedly different. During the growth phase, there would be a permanent presence of unsatisfied open domains at both ends of each growing fibril, providing sticky ends for continuing growth. Prior to the formation of stable seeds, however, the unfolded monomers (presumed to be in equilibrium with the folded form) could always refold back as monomers or closed-ended circular oligomers, thus eliminating the points of attachment (unsatisfied domains) from the system. Here formation of closed oligomers (e.g., dimers) could be a dead-end on the oligomerization pathway, and in a way could prevent, or slow down, the process of amyloid fibril formation.[17]

It is also conceivable that in amyloidogenic aggregation two different mechanisms could be in operation. The building blocks of the amyloid fibril could be formed as closed-ended domain-swapped oligomers, which would aggregate further using a different mechanism, for example, intermolecular β-type association. In this scenario, the accumulation of closed-ended oligomers would not be viewed as a suicidal trap on the polymerization pathway[17] but would be a required initial step providing "substrates" for the final polymerization step. The present structure is suggestive of this possibility. It shows that domain-swapped dimers of human cystatin C can associate further through large intermolecular β-structures, created by joining the free edges of the molecular β-sheets. In the previously reported structures of 3D domain-swapped cystatin C, extensive β-type aggregation was hindered because the β-chains in the two folding units of the dimer were roughly perpendicular. In contrast, in the present structure, they are aligned in one direction, thus facilitating intermolecular β-type association, such as seen in the perfect β5–β5 interactions that exploit the dimeric nature of the interacting units.

The interactions at the other, β2, edge of the molecular β-sheet are less perfect and involve only side-chain, but not main-chain, hydrogen bonding. On the other hand, they provoke speculations on the significance of the flexible N-terminal segment, which could play a protective role, blocking intermolecular access to strand β2, which (considering the short length of the β1 segment) becomes the effective edge of the molecular β-sheet. This speculation is interesting, considering the fact that the protein material extracted from cystatin C amyloid fibrils is truncated at the N-terminus.[15]

CONCLUSIONS

The accumulating structural data,[12,16] evidence from protein engineering experiments,[11] and the fact that HCC dimers are observed in the cerebrospinal fluid and blood

plasma of HCCAA patients suffering from cystatin C amyloidosis,[9] all suggest that aggregation of human cystatin C into amyloid fibrils very likely involves 3D domain swapping. However, the precise molecular mechanism of this process is still not clear. The present crystal structure, in which 3D domain-swapped HCC dimers form tetrameric units using strong β-sheet interactions, and then associate further to produce a perpendicular alignment of all β-strands relative to a common direction, is suggestive of the possibility that in HCC fibril formation two processes could be involved. In a speculative scenario, 3D domain swapping would provide dimeric building blocks, which would associate further using extended β interactions at the free edges of the molecular β-sheets.

The present structure is also important because it demonstrates that dimeric cystatin C is very flexible and can assume markedly different conformations by modulations of the geometry of the hinge region at the open interface, while preserving, with very high degree of precision, the monomeric fold recreated through the closed interface. This reinforces the notion that molecular intervention targeting the L1 sequence of HCC, such as in the preliminary experiments using inactive papain,[11] might be a promising therapeutic approach for the treatment of cystatin C–related amyloidoses. Moreover, the results of the present study also suggest that similar therapeutic intervention could use specific antibodies targeting the oligopeptide sequences forming the β2 and β5 edges of the cystatin C β-sheet structure.

ACKNOWLEDGMENTS

Some of the calculations were done at the Poznan Metropolitan Supercomputing and Networking Center.

REFERENCES

1. Olafsson I, Grubb A. Hereditary cystatin C amyloid angiopathy. Amyloid 2000;7:70–79.
2. Maruyama K, Ikeda S, Ishihara T, Allsop D, Yanagisawa N. Immunohistochemical characterization of cerebrovascular amyloid in 46 autopsied cases using antibodies to beta-protein and cystatin C. Stroke 1990;21:397–403.
3. Grubb A, Löfberg H. Human gamma-trace, a basic microprotein: amino acid sequence and presence in the adenohypophysis. Proc Natl Acad Sci USA 1982;79:3024–3027.
4. Grubb A, Löfberg H, Barrett AJ. The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin. FEBS Lett 1984; 170:370–374.
5. Abrahamson M, Alvarez-Fernandez M, Nathanson C-M. Cystatins. Biochem Soc Symp 2003;70:179–199.
6. Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. The 2.0 Å X-ray crystal-structure of chicken egg-white cystatin and its possible mode of interaction with cysteine proteinases. EMBO J 1988;7:2593–2599.
7. Alvarez-Fernandez M, Barrett AJ, Gerhartz B, Dando PM, Ni J, Abrahamson M. Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. J Biol Chem 1999;274:19195–19203.
8. Abrahamson M, Grubb A. Increased body temperature accelerates aggregation of the (Leu-68→Gln) mutant cystatin C, the amyloid-forming protein in hereditary cystatin C amyloid angiopathy. Proc Natl Acad Sci USA 1994;91:1416–1420.
9. Bjarnadottir M, Nilsson C, Lindström V, Westman A, Davidsson P, Thormodsson F, Blöndal H, Gudmundsson G, Grubb A. The cerebral hemorrhage-producing cystatin C variant (L68Q) in extracellular fluids. Amyloid 2001;8:1–10.
10. Ekiel I, Abrahamson M. Folding related dimerization of human cystatin C. J Biol Chem 1996;271:1314–1321.
11. Nilsson M, Wang X, Rodziewicz-Motowidlo S, Janowski R, Lindström V, Önnerfjord P, Westermark G, Grzonka G, Jaskolski M, Grubb A. Prevention of domain swapping inhibits dimerization and amyloid fibril formation of cystatin C: use of engineered disulfide bridges, antibodies and carboxymethylpapain to stabilize the monomeric form of cystatin C. J Biol Chem 2004;279: 24236–24245.
12. Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Alvarez Fernandez M, Abrahamson M, Jaskolski M. Human cystatin C—a protein with amyloidogenic properties dimerizes through domain swapping. Nat Struct Biol 2001;8:316–320.
13. Bennett MJ, Choe S, Eisenberg DS. Domain swapping—entangling alliances between proteins. Proc Natl Acad Sci USA 1994;91: 3127–3131.
14. Bennett MJ, Schlunegger MP, Eisenberg D. 3D domain swapping—a mechanism for oligomer assembly. Protein Sci 1995;4: 2455–2468.
15. Ghiso J, Jensson O, Frangione B. Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C). Proc Natl Acad Sci USA 1986;83:2974–2978.
16. Janowski R, Abrahamson M, Grubb A, Jaskolski M. Domain swapping in N-truncated human cystatin C. J Mol Biol 2004;341: 151–160.
17. Jaskolski M. 3D domain swapping, protein oligomerization, and amyloid formation. Acta Biochim Pol 2001;48:807–827.
18. Knaus KJ, Morillas M, Swietnicki W, Malone M, Surewicz WK, Yee VC. Crystal structure of the human prion protein reveals a mechanism for oligomerization. Nat Struct Biol 2001;8:770–774.
19. Lee S, Eisenberg D. Seeded conversion of recombinant prion protein to a disulfide-bonded oligomer by a reduction–oxidation process. Nat Struct Biol 2003;10:725–730.
20. Glenner GG. Amyloid deposits and amyloidosis—the β-fibrilloses 1. N Engl J Med 1980;302:1283–1292.
21. Glenner GG. Amyloid deposits and amyloidosis—the β-fibrilloses 2. N Engl J Med 1980;302:1333–1343.
22. Spinelli S, Desmyter A, Frenken L, Verrips T, Tegoni M. Cambillau Ch. Domain swapping of a llama VHH domain builds a crystal-wide β-sheet structure. FEBS Lett 2004;564:35–40.
23. Kozak M, Jankowska E, Janowski R, Grzonka Z, Grubb A, Fernandez MA, Abrahamson M, Jaskolski M. Expression of a selenomethionyl derivative and preliminary crystallographic studies of human cystatin C. Acta Crystallogr D Biol Crystallogr 1999;55:1939–1942.
24. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
25. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. J Appl Crystallogr 1997;30:1022–1025.
26. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR System: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
27. Jones T, Zou, J, Cowan S, Kjeldgaard M. Improved methods for building protein models in electron density maps and location of errors in these models. Acta Crystallogr A 1991;47:110–119.
28. Matthews BW. Solvent content of protein crystals. J Mol Biol 1968;33:491–497.
29. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
30. Koradi R, Billeter M, Wuthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graphics 1996;14:51–55.
31. Engh RA, Dieckmann T, Bode W, Auerswald EA, Turk V, Huber R, Oschkinat H. Conformational variability of chicken cystatin—comparison of structures determined by X-ray-diffraction and NMR-spectroscopy. J Mol Biol 1993;234:1060–1069.
32. Schlunegger MP, Bennett MJ, Eisenberg D. Oligomer formation by 3D domain swapping: a model for protein assembly and misassembly. Adv Protein Chem 1997;50:61–122.
33. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876–4882.

34. Cohen GH. ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. J Appl Crystallogr 1997;30:1160–1161.
35. Staniforth RA, Giannini S, Higgins LD, Conroy MJ, Hounslow AM, Jerala R, Craven CJ, Waltho JP. Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. EMBO J 2001;20:4774–4781.
36. CueMol—Molecular visualization framework. Available online at http://cuemol.sourceforge.jp/en/
37. ViewerLite v 4.2, Accelrys Inc. Available online at http://www.accelrys.com
38. Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CCF. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. J Mol Biol 1997;273:729–739.
39. Wille H, Michelitsch MD, Guenebaut V, Supattapone S, Serban A, Cohen FE, Agard DA, Prusiner SB. Structural studies of the scrapie prion protein by electron crystallography. Proc Natl Acad Sci USA 2002;99:3563–3568.
40. Rawlings ND, Tolle DP, Barrett AJ. MEROPS: the peptidase database. Nucleic Acids Res 2004;32(Database issue);D160–D164. Available online at http://merops.sanger.ac.uk
41. Newcomer ME. Protein folding and three-dimensional domain swapping: a strained relationship? Curr Opin Struct Biol 2004;12:48–53.
42. Ogihara NL, Ghirlanda G, Boryson JW, Gingery M, DeGrado WF, Eisenberg D. Design of three-dimensional domain-swapped dimers and fibrous oligomers. Proc Natl Acad Sci USA 2001;98:1404–1409.
43. Liu Y, Gotte G, Libonati M, Eisenberg D. A domain-swapped RNase A dimer with implications for amyloid formation. Nat Struct Biol 2001;8:211–214.
44. Liu Y, Gotte G, Libonati M, Eisenberg D. Structures of the two 3D domain-swapped RNase A trimers. Protein Sci 2002;11:371—380.