



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: [www.elsevier.com/locate/ybbrc](http://www.elsevier.com/locate/ybbrc)



## β2-Strand of salivary S cystatins: A “chameleon sequence”

Alberto Vitali<sup>a</sup>, Cristiana Carelli Alinovi<sup>a</sup>, Maria Cristina De Rosa<sup>a</sup>, Raffaele Petruzzelli<sup>b,\*</sup>

<sup>a</sup> Istituto di Biochimica e Biochimica Clinica, Facoltà di Medicina, Università Cattolica e/o Istituto per la Chimica del Riconoscimento Molecolare, C.N.R., Largo F. Vito 1, 00168 Roma, Italy

<sup>b</sup> Dipartimento di Scienze Biomediche, Università degli Studi “G. D’Annunzio”, Via dei Vestini 29, 66013 Chieti, Italy

### ARTICLE INFO

#### Article history:

Received 17 June 2009

Available online 23 June 2009

#### Keywords:

Salivary cystatins

β-Strand

α-Helix

CD

Molecular modeling

Folding

Peptide

Capping box

### ABSTRACT

Secondary structure prediction of salivary cystatins S, SA, and SN carried out by several methods label the 39–58 sequence (β2-strand) as predominantly α-helical. The helical propensity of a peptide corresponding to β2-strand of salivary S cystatin analyzed by CD display high helical propensity in aqueous solution, whereas peptides matching the β2-strand amino acid sequence of cystatins S and SN, display random coil conformation in aqueous solution but acquire α-helical conformation in the presence of trifluoroethanol (TFE). Moreover molecular dynamics simulation performed on the homology modeling of cystatin SA constructed on the basis of recently determined three-dimensional structure of salivary cystatin D, suggests that cystatin SA does not significantly deviate from the starting structure over the course of the simulation. The results obtained indicate that the β2-strand of salivary S cystatins has high helical propensity when isolated from native protein and acquire the final β structure by interaction with the rest of the polypeptide chain.

© 2009 Elsevier Inc. All rights reserved.

### Introduction

The cystatins superfamily is comprised of a large group of related proteins, most of which are potent inhibitors of cysteine peptidases. The mammalian superfamily members are of a three major types [1,2].

Type-1 cystatins (also called stefins) are primarily cytoplasmatic single-domain proteins composed of about 100 amino acids residues, with no disulfide bridges and no signal peptide. Type-2 cystatins are secreted inhibitors and are also single-proteins but are about 120 amino acid residues long and present two well-conserved disulfide bridges and typical signal peptides. Type-3 cystatins or kininogens, are multidomain proteins presenting three tandemly repeated type-2 cystatin-like domains.

In human saliva, type-2 cystatins are secreted from submandibular and sublingual glands namely S [3], SN [4], and SA [5] and one cystatin D [6] secreted from the parotid glands. Various functions have been ascribed to these proteins: (a) protective action against proteases of endogenous origin (secreted by tissues) as a result of tissues necrosis and inflammation or against proteases secreted by exogenous agents; (b) modulation of the immuno system; (c) antibacterial and antiviral activities (unrelated to inhibition of cysteine-protease) and a role in control of mineralization at the tooth surface.

The three-dimensional structure of type-1 (human stefin A and B) and type-2 cystatins (Chicken egg-white, cystatin D) have been determined by NMR [7] and X-ray crystallography [8,9]. All these cystatins show the same overall structure, with a five-stranded antiparallel β-sheet wrapped around a five-turn α-helix.

Despite the secondary structural elements are strictly conserved in all the known cystatins three-dimensional structures, several secondary structure predictive methods evidenced that the segments (cystatin SA numbering 39–58) corresponding to β2-strand of the three salivary S cystatins adopt an α-helix conformation.

The acquisition of different conformations by the same polypeptide sequence has been evidenced by several authors and in particular Minor and Kim [10] have designed a “chameleon” peptide that folds as α-helix or a β-strand when placed in different regions of the sequence of the immunoglobulin-binding domain of protein G.

An important consequence of this behavior is that a shift to non-native conformations may trigger the development of amyloidogenic proteins with the onset of serious diseases.

To analyze the structural propensity of this region isolated from the native proteins, we synthesized several peptides corresponding to the β2-strand of the salivary cystatins S, SA, and SN and their structural propensity in water and TFE/water solutions has been evaluated by CD. Furthermore, to observe regional backbone conformational changes and β2-strand stability, molecular dynamics simulation has been applied to the homology model of salivary

\* Corresponding author. Fax: +39 0871 3554736.

E-mail address: [r.petruzzelli@dsb.unich.it](mailto:r.petruzzelli@dsb.unich.it) (R. Petruzzelli).

SA cystatin constructed on the basis of recently determined three-dimensional structure of salivary cystatin D [9].

The results obtained, indicate that we are in the presence of a sequence that when isolated from the native protein has high helical propensity which is not related to the cystatin secondary structure in the native state, and suggest a mechanism of non-hierarchical protein folding.

## Materials and methods

**Sequence analysis and secondary structure prediction for cystatins.** Predictions were carried out at the ExPASy proteomics server site (<http://us.expasy.org/>). Methods selected were GOR [11], SOPMA [12], Jpred [13], and HNN [14].

**Peptide synthesis and design.** The following synthetic peptides matching the amino acid sequence corresponding to the fragment 39–58 of salivary cystatins S, SA, and SN and comprising the loop 1 and the  $\beta$ 2-strand, were purchased from Peptide Specialty Laboratories GmbH (Germany).

Name	Sequence
SA	TEDEYYRRLRLVLRAREQIV
S	TEDEYYRRPLQLVLRARQETF
SN	TKDDYYRRPLRLVLRARQQTVA

The residues in bold correspond to the capping box.

Moreover, were also synthesized a peptide S (TEDEYYRR LLQVLRARQETF) with the substitution **P9L** and a peptide SA (TED**A**YYRRLRLVLRAREQIV) with the substitution **E4A**.

All the peptides were purified by HPLC and analyzed by mass spectrometry. Their purity was >98%.

**CD analysis.** CD spectra were obtained on a Jasco J-600 spectrophotometer equipped with a thermostatic temperature controller. CD spectra were recorded in quartz cell of 0.1 cm of pathlength at 25 °C between 190 and 250 nm, using a 2.0 nm bandwidth and a scanning rate of 20 nm/min with a wavelength step of 0.1 nm and a time constant of 0.1 s. All the peptides at the concentration 80  $\mu$ M were dissolved in 10 mM sodium phosphate buffers at pH 7.0. Different TFE/aqueous buffer mixtures were employed. CD bands intensities are expressed as molar ellipticities,  $([\theta])_M$  in  $\text{deg cm}^2 \text{dmol}^{-1} \times 10^{-3}$ .

**Homology modeling of human cystatin SA.** The homology model of human cystatin SA was built using MODELLER, v. 7 [15] as implemented in InsightII (Accelrys Inc.) based on the crystal structure of human cystatin D (pdb code: 1RN7 [9]). A pairwise alignment between human cystatin SA and human cystatin D using Clustal W revealed a sequence identity of 54%. A collection of 20 model structures was generated and ranked based on an objective function score provided by MODELLER. From this ensemble, the model with the lowest objective function was checked with Procheck [16] and chosen as the representative model for further study. The coordinates of this model was then subjected to the Verify3D algorithm [17] (using the Verify3D Structure Evaluation Server (available at: [nihserver.mbi.ucla.edu/Verify\\_3D](http://nihserver.mbi.ucla.edu/Verify_3D)) to identify regions of improper folding. Disulfide bonds between Cys75–Cys85 and Cys99–Cys119 of 1RN7 template structure were constrained during the modeling.

**Molecular dynamics simulation of cystatin SA model.** To probe the possibility of structural changes in the three-dimensional structure, molecular dynamics simulation of cystatin SA model was performed. The protein was immersed in a rectangular box filled with TIP3 water molecules. Periodic boundary conditions (PBC) were employed to avoid edge effects. All simulations were

performed with the CHARMM program, version c31b1 [18] as implemented in InsightII (Accelrys Inc.) using the parameter set from the CHARMM22 force field and a constant dielectric  $\epsilon = 1$ . After solvation, the simulation system was energy-minimized with 200 iterations of the steepest descent algorithm followed by MD for 1 ns, with the first 12 ps taken for heating and the following 12 ps for equilibration. In all energy evaluations, a 10 Å non-bonded cutoff distance was employed, with a shifting function applied between 8 and 10 Å for van der Waals and electrostatic terms to eliminate discontinuities due to the cutoff. The Verlet algorithm was used with a 2-fs time step and SHAKE constraints for all bonds of hydrogens to heavy atoms. All computations were performed on an IBM Intellistation running Red Hat Linux 3.0. Analysis of MD trajectories was performed with the VMD 1.8.5 package and associated plug-ins [19]. The root mean square deviation of the complexes from the corresponding initial structure was calculated over the MD trajectory. The secondary structure of the protein along the trajectory was calculated using STRIDE [20] implemented in VMD.

## Results

### Secondary structure prediction for S cystatins

A secondary structure prediction was run for salivary cystatin S, SA, and SN using the ExPASy server as described in Materials and methods. The overall prediction of  $\beta$ -strand fits nicely with experimental data, except for a region comprising the  $\beta$ 2-strand where there is ambiguity in the prediction among the several methods tested.

### CD analysis of the designed peptides

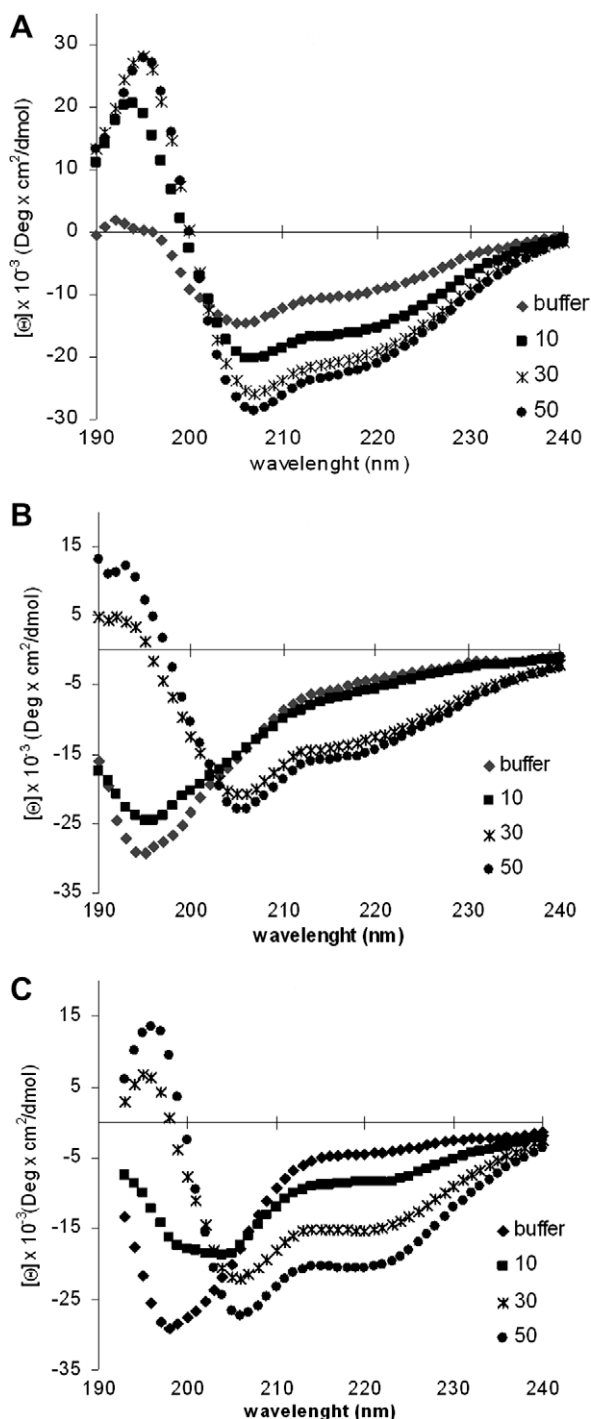
To analyze the propensity of the fragments corresponding to the  $\beta$ 2-strand isolated from the protein, we synthesized several peptides corresponding to the 39–58 segment of cystatins S, SA, and SN. An  $\alpha$ -helical prediction on isolated peptides performed by AGADIR algorithm [21] predicts for  $\beta$ 2-strand of cystatin SA an average helicity of 51% in water at pH 7 and 25 °C. All the peptides showed no tendency to aggregate in aqueous solution and their CD spectra in water and water/TFE are shown in Fig. 1.

The far-UV CD spectra show that the fragments corresponding to cystatins S and SN are predominantly in random coil conformation, while fragment SA showed in the same conditions negative signals at 208 and 222 nm indicating the presence of an  $\alpha$ -helix.

To better investigate the helix propensity of these peptides, we have performed CD spectra in the presence of increasing TFE concentrations. Upon addition of TFE, the spectra of the peptides S, SN changed to that of an  $\alpha$ -helix. A highly cooperative coil-to-helix transition, with an isodichroic point at 203 nm was seen for the three peptides. In TFE/water, the peptide SA undergoes an increase of the ellipticity showing a further stabilization of the helical structure. This fragment showed the higher helical content among the other tested peptide reaching an ellipticity percentage of about 84% ( $-30 \times 10^{-3} \text{ deg cm}^2 \text{dmol}^{-1}$ ).

These results indicate that peptide SA has high helical propensity in water, as predicted by the Agadir algorithm, while fragments derived from  $\beta$ 2-strand region of salivary cystatins S and SN exhibited high helical propensity only in the presence of TFE. Moreover, these results, indicate that the helical content of a peptide in TFE is more related to the helical content predicted by the secondary structure prediction methods rather than that of the native state.

To investigate the high helical propensity exhibited by the peptide SA in water with respect to the other fragments, we performed



**Fig. 1.** Far-UV CD spectra of the peptides SA (A), S (B), and SN (C). The spectra were recorded at 80  $\mu$ M peptide concentration in aqueous buffer and in solutions of TFE/buffer at different TFE concentrations (see figure inset legends).

a comparison of the amino acid sequence of the three peptides and evidenced the presence of a leucine residue in the SA peptide substituted by a proline in the S and SN peptides.

In order to demonstrate the role of this proline residue in lowering the helical propensity of the fragments derived from the cystatins S and SN, a peptide matching the amino acid sequence of the  $\beta$ 2-strand of cystatin S with a substitution P/L was synthesized. The far-UV CD spectra of this peptide in water changed from random coil to  $\alpha$ -helix and the addition of TFE induced an enhancing of helicity.

Moreover, the salivary cystatins S, SA and SN present at the beginning of the strand  $\beta$ 2, a structural motif constituted by the amino sequence TXXD/E known as “Capping box” [22–24]. This motif has been suggested to exert a strong influence on the nucleation and folding of the helix N-terminus [25]. To demonstrate the influence of this motif in inducing the strong helix propensity, we synthesized an SA peptide with the substitution E4/A. This peptide was studied by CD in water and water/TFE. The peptide in water maintains its  $\alpha$ -helical structure, but the addition of increasing concentrations of TFE do not increase its helicity. The results obtained, confirm that all the peptides corresponding to the  $\beta$ 2-strand have an intrinsic propensity to adopt an  $\alpha$ -helical structure enhanced by the presence of a capping box motif at their N-terminus.

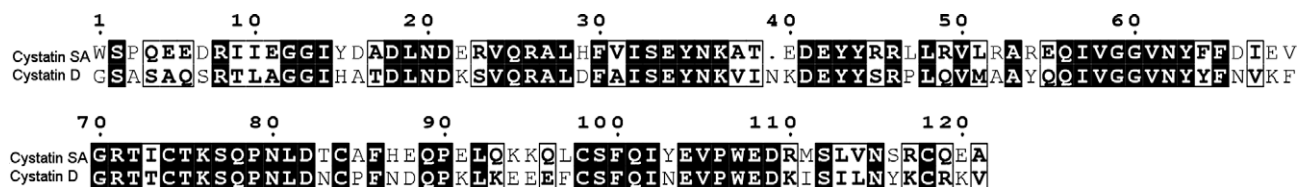
#### Homology modeling

In order to verify the correct folding of the  $\beta$ 2-strand in the native protein and its stability, we decided to construct a three-dimensional model of the SA cystatin and perform a molecular dynamics simulation. Previous homology model of human cystatin SA was based on chicken cystatin structure [8]. With the availability of the three-dimensional structure of human cystatin D [9] a more accurate model was constructed because of the higher sequence homology (sequence identity of 54% versus 40%). The first stage of modeling was performed with the program MODELLER using 1RN7 pdb structure and the alignment from Clustal W as input (Fig. 2). Twenty models were generated by the program MODELLER which implements an automated approach to comparative protein structure modeling by satisfaction of spatial restraints. The C $\alpha$ -chain trace of the 20 models of cystatin SA superimposed to the template structure shows negligible positional variability between the models. Among the obtained conformations one was selected on the basis of the lowest MODELLER objective function. The quality of the emerging model was monitored by Procheck. 94% of modeled residues fall within the most favored regions of the Ramachandran plot as defined by Procheck, with no residues in disallowed regions. The goodness of the selected folding was assessed by Verify3D which evaluates the compatibility of a given residue in a certain three-dimensional environment. A score below zero for a given residue means that the conformation adopted by that residue in the model is not compatible with its surrounding environment. The 3D–1D Verify3D scores presented by our model are always positive and similar to those presented by the template structure. Despite the absence of negative score regions in the plots, it is worth noticing that the region between residues 39–58, which corresponds to  $\beta$ 2-strand, displays the lowest 3D–1D scores. Fig. 3 shows ribbon diagram for cystatin SA homology-built model. The model structure presents the typical cystatin fold, with a five-stranded antiparallel  $\beta$ -sheet wrapped around a five-turn  $\alpha$ -helix.

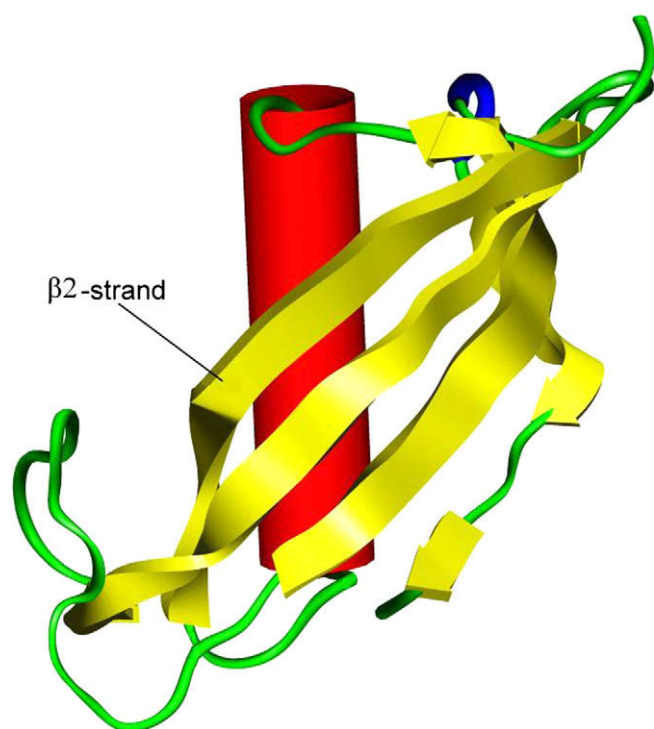
#### Molecular dynamics simulation

To consider a three-dimensional structural change in aqueous environment, the homologous three-dimensional model was refined by MD simulation. MD simulation was performed on the solvated homology model cystatin SA for a time scale of 1 ns under constant pressure and temperature conditions.

MD trajectory equilibrated within 200 ps and analysis of MD trajectories is reported in Fig. 4. As an indication of the degree of structural drift during simulation, the positional root mean square deviation (RMSD) from the starting structure was calculated as a function of the simulation time for all C $\alpha$  atoms (Fig. 4A). As shown in the figure, over the whole 1-ns trajectory, the structural change was moderate with a C $\alpha$  RMSD of 1.7 Å at the end of the simula-



**Fig. 2.** Sequence alignment of human cystatin SA with human cystatin D. Sequences were aligned using Clustal W and modified for print representation using ESPrnt 2.2, available at [esprnt.ibcp.fr/ESPrnt/ESPrnt/index.php](http://esprnt.ibcp.fr/ESPrnt/ESPrnt/index.php) [37]. The residues in black boxes are identical in the homologs, whereas the residues in white boxes are similar.



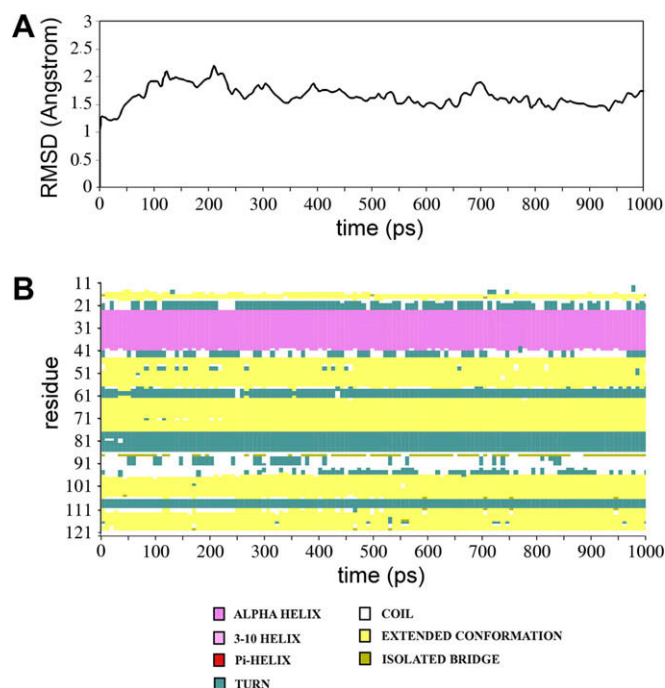
**Fig. 3.** Molecular model of cystatin SA. Protein structure is shown as a ribbon diagram.  $\beta$ -Sheets are shown in yellow,  $\alpha$ -helices in red, turns in blue and loops in green.

tion. On the whole, the above analysis suggests that cystatin A does not significantly deviate from the starting structure over the course of the simulation. Analysis of the MD trajectories in terms of secondary structure elements does enable a local structural analysis complementing the above characterization of the C $\alpha$  conformational dynamics. The time evolution of the secondary structure elements is depicted in Fig. 4B and displays that the secondary structure is well preserved over the MD trajectory. It is worth noticing that the  $\beta$ 2-strand shows a slight instability which is for instance absent in  $\beta$ 3-strand.

## Discussion

Several mechanisms have been proposed to explain the efficient folding of globular proteins, the most popular of which invokes the existence of a molten globule intermediate and sequential and hierarchical folding [26,27]. In the nucleation–condensation mechanism the formation of the nucleus and of a secondary structure is a concerted event for small proteins [28–31].

Whereas the hierarchical model has been successful in explaining the folding of several all- $\alpha$  proteins, it does not provide a satisfactory explanation of the folding of all- $\beta$  proteins, where local formation of a  $\beta$ -strand depends on long-range interactions with another preformed  $\beta$ -strand.



**Fig. 4.** Analysis of MD trajectories. Conformational changes compared with the starting structure are measured by the root mean square deviation (RMSD) of C $\alpha$  atoms during the MD run (A). Classification of MD trajectory in terms of secondary structure elements was generated by STRIDE in VMD (B).

The three-dimensional structure of the cystatins is formed by five-stranded antiparallel  $\beta$ -sheets wrapped around an  $\alpha$ -helix. The  $\beta$ 2-strand of this  $\beta$ -sheet is predicted to be an  $\alpha$ -helix, and this propensity was confirmed by spectroscopic analysis of some peptides matching the sequence of  $\beta$ 2-strand of salivary cystatins.

The results presented in this manuscript suggest that, although the local interactions favor the formation of an  $\alpha$ -helix, long-range interactions predominate in the native state. This situation is in accord with the model of non-hierarchical folding described by several groups [32,33], which propose that folding is initiated on the basis of local “short-range” interactions that are subsequently disrupted or rearranged in order for the protein to fold to the native  $\beta$ -sheet structure.

This folding strategy is common to several  $\beta$ -protein, in fact Sanz et al. [34] have analyzed by CD and nuclear magnetic resonance (NMR) the helical propensity of all  $\beta$ -protein, acidic fibroblast growth factor and several peptides corresponding to strands  $\beta$ 8 and  $\beta$ 9. They found that the region corresponding to  $\beta$ 8/9 peptide has high helical propensity suggesting a mechanism of non-hierarchical protein folding.

Similar results have been obtained for the N-terminal  $\beta$ -hairpin fragment of the streptococcal protein G B1 domain [35] and for several peptides fragments derived from  $\beta$ -lactoglobulin [36] a predominantly  $\beta$ -sheet protein, which exhibited marked helical

propensity in TFE. It is important to underline, that while the presence at the N-terminal of  $\beta$ 2-strand of a capping box confers to the isolated fragments high helical propensity, the same motif does not influence the correct formation of the  $\beta$ 2-strand in the native structure. This motif may be important in the first folding events, in inducing a local formation of an  $\alpha$ -helix, successively lost in favor of long-range interactions. In fact, the analysis of the three-dimensional structure of both cystatin D and cystatin SA model, shows, that the side chains of the residues constituting the capping box do not form the characteristic hydrogen bonding pattern, but are exposed on the surface of the protein.

A 1.0-ns molecular dynamics simulation (MD) is certainly not long enough to observe folding or unfolding of a protein. However, it is long enough to observe regional backbone conformational changes involving a partial unfolding of a 3D structure that was incorrectly folded by homology modeling. The MD results clearly indicate that the secondary structure is well conserved over the whole trajectory. In conclusion, our combined experimental and computational investigations, indicate that the intrinsic helical propensity of the peptide fragments derived from the salivary cystatins is not related to their secondary structure in the native state, and suggest a case of non-hierarchical protein folding mechanism.

## Acknowledgments

The authors acknowledge the financial support of Chieti University and of Italian MIUR according to their programs of scientific research promotion and diffusion.

## References

- [1] M. Abrahamson, M. Alvarez-Fernandez, C.M. Nathanson, Cystatins, *Biochem. Soc. Symp.* 70 (2003) 179–199.
- [2] N.D. Rawlings, A.J. Barrett, Evolution of proteins of the cystatin superfamily, *J. Mol. Evol.* 30 (1990) 60–71.
- [3] S. Isemura, E. Saitoh, S. Ito, M. Isemura, K. Sanada, Cystatin S: a cysteine proteinase inhibitor of human saliva, *J. Biochem.* 96 (1984) 1311–1314.
- [4] S. Isemura, E. Saitoh, K. Sanada, Characterization of a new cysteine proteinase inhibitor of human saliva, cystatin SN, which is immunologically related to cystatin S, *FEBS Lett.* 198 (1986) 145–149.
- [5] S. Isemura, E. Saitoh, K. Sanada, Characterization and amino acid sequence of a new acidic cysteine proteinase inhibitor (cystatin SA) structurally closely related to cystatin S, from human whole saliva, *J. Biochem.* 102 (1987) 693–704.
- [6] J.P. Freije, M. Abrahamson, I. Olafsson, G. Velasco, A. Grubb, C. López-Otín, Structure and expression of the gene encoding cystatin D, a novel human cysteine proteinase inhibitor, *J. Biol. Chem.* 266 (1991) 20538–20543.
- [7] J.R. Martin, R. Jerala, L. Kroon-Zitko, E. Zervovnik, V. Turk, J.P. Waltho, Structural characterisation of human stefin A in solution and implications for binding to cysteine proteinases, *Eur. J. Biochem.* 225 (1994) 1181–1194.
- [8] W. Bode, R. Engh, D. Musil, U. Thiele, R. Huber, A. Karshikov, J. Brzin, J. Kos, V. Turk, The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases, *EMBO J.* 7 (1988) 2593–2599.
- [9] M. Alvarez-Fernandez, Y.H. Liang, M. Abrahamson, X.D. Su, Crystal structure of human cystatin D, a cysteine peptidase with restricted inhibition profile, *J. Biol. Chem.* 280 (2005) 18221–18228.
- [10] D.L. Minor, P.S. Kim, Context-dependent secondary structure formation of a designed protein sequence, *Nature* 380 (1996) 730–734.
- [11] J. Garnier, D.J. Osguthorpe, B. Robson, Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins, *J. Mol. Biol.* 120 (1978) 97–120.
- [12] C. Geourjon, G. Deléage, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, *Comput. Appl. Biosci.* 11 (1995) 681–684.
- [13] J.A. Cuff, M.E. Clamp, A.S. Siddiqui, M. Finlay, G.J. Barton, JPred: a consensus secondary structure prediction server, *Bioinformatics* 14 (1998) 892–893.
- [14] Y. Guermeur, Combinaison de classificateurs statistiques, Application à la prédiction de structure secondaire des protéines, Ph.D. thesis.
- [15] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [16] R.A. Laskowski, J.A. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR* 8 (1996) 477–486.
- [17] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature* 356 (1992) 83–85.
- [18] B. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM a program for macromolecular energy, minimisation and dynamics calculation, *J. Comput. Chem.* 4 (1983) 187–217.
- [19] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph. Model.* 14 (1996) 33–38.
- [20] D. Frishman, P. Argos, Knowledge-based protein secondary structure assignment, *Proteins* 23 (1995) 566–579.
- [21] V. Munoz, L. Serrano, Development of the multiple sequence approximation within the Agadir model of  $\alpha$ -helix formation. Comparison with Zimm–Bragg and Lifson–Roig Formalisms, *Biopolymers* 41 (1997) 495–509.
- [22] E.N. Baker, R.E. Hubbard, Hydrogen bonding in globular proteins, *Prog. Biophys. Mol. Biol.* 44 (1984) 97–179.
- [23] S. Dasgupta, J.A. Bell, Design of helix ends. Amino acid preferences, hydrogen bonding and electrostatic interactions, *Int. J. Pept. Protein Res.* 41 (1993) 499–511.
- [24] E.T. Harper, G.D. Rose, Helix stop signals in proteins and peptides: the capping box, *Biochemistry* 32 (1993) 7605–7609.
- [25] M.A. Jimenez, Rico M. Munoz, L. Serrano, Helix stop and start signals in peptides and proteins. The capping box does not necessarily prevent helix elongation, *J. Mol. Biol.* 242 (1994) 487–496.
- [26] K. Kuwajima, The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure, *Proteins: Struct. Funct. Genet.* 6 (1989) 87–103.
- [27] O.B. Ptitsyn, The molten globule state, in: T.E. Creighton (Ed.), *Protein Folding*, WH Freeman and Company, New York, 1992, pp. 243–300.
- [28] A.R. Fersht, Optimization of the rates of protein folding: the nucleation–condensation mechanism and its implications, *Proc. Natl. Acad. Sci. USA* 92 (1995) 10869–10873.
- [29] E.I. Shakhovich, Modeling protein folding: the beauty and power of simplicity, *Folding Des.* 1 (1996) R50–R54.
- [30] R.L. Baldwin, The nature of the protein folding pathways: the classical versus the new view, *J. Biochem. NMR* 5 (1995) 103–109.
- [31] L.S. Itzhaki, D.E. Otzen, A.R. Fersht, The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation–condensation mechanism for protein folding, *J. Mol. Biol.* 254 (1995) 260–288.
- [32] K. Shiraki, K. Nishikawa, Y. Goto, Trifluoroethanol-induced stabilization of the alpha-helical structure of beta-lactoglobulin: implication for non-hierarchical protein folding, *J. Mol. Biol.* 245 (1995) 180–194.
- [33] G. Chikenji, M. Kikuchi, What is the role of non-native intermediates of beta-lactoglobulin in protein folding?, *Proc. Natl. Acad. Sci. USA* 97 (2000) 14273–14277.
- [34] J.M. Sanz, N.A. Jimenez, G. Gimenez-Gallego, Hints of nonhierarchical folding of acidic fibroblast growth factor, *Biochemistry* 41 (2002) 1923–1933.
- [35] F.J. Blanco, L. Serrano, Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements, *Eur. J. Biochem.* 230 (1995) 634–649.
- [36] D. Hamada, Y. Kuroda, T. Tanaka, Y. Goto, High helical propensity of the peptide fragments derived from  $\beta$ -lactoglobulin, a predominantly  $\beta$ -sheet protein, *J. Mol. Biol.* 254 (1995) 737–746.
- [37] D.I. Stuart, F. Metoz, ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics* 15 (1999) 305–308.