Universidade de Aveiro Departamento de Biologia
2007

Cláudia Sofia Soares de Oliveira

Estados conformacionais da cardosina A

Universidade de Aveiro Departamento de Biologia
2007

Cláudia Sofia Soares de Oliveira

Estados conformacionais da cardosina A

tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Biologia, realizada sob a orientação científica da Professora Doutora Marlene Maria Tourais de Barros, Professora Associada do Centro Regional das Beiras, Universidade Católica Portuguesa e do Professor Doutor Henrique Manuel Apolónia Coutinho Fonseca, Professor Auxiliar do Departamento de Biologia da Universidade de Aveiro.

Apoio financeiro do POCTI no âmbito do III Quadro Comunitário de Apoio através do projecto POCTI/QUI/60791/2004

Apoio financeiro da Universidade de Aveiro através de bolsa de Doutoramento

o júri

presidente

Professor Doutor Francisco Cardoso Vaz
Vice Reitor da Universidade de Aveiro

Professor Doutor Amadeu Mortágua Velho da Maia Soares
Professor Catedrático do Departamento de Biologia da Universidade de Aveiro

Professora Doutora Maria Ana Dias Monteiro Santos
Professora Catedrática do Departamento de Biologia da Universidade de Aveiro

Professor Doutor Euclides Manuel Vieira Pires
Professor Associado da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Professor Doutor Carlos José Fialho da Costa Faro
Professor Associado da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Professor Doutor António Carlos Matias Correia
Professor Associado com Agregação do Departamento de Biologia da Universidade de Aveiro

Professora Doutora Marlene Maria Tourais de Barros
Professora Associada do Centro Regional das Beiras da Universidade Católica Portuguesa

Professor Doutor Henrique Manuel Apolónia Coutinho Fonseca
Professor Auxiliar do Departamento de Biologia da Universidade de Aveiro

agradecimentos

Agradeço à Universidade de Aveiro a bolsa de estudos concedida e pela confiança transmitida ao longo deste anos. Agradeço também ao Departamento de Biologia da Universidade de Aveiro, bem como ao Centro de Biologia Celular por possibilitar a realização do trabalho aqui apresentado.

À Professora Doutora Marlene Barros, orientadora deste projecto, pelo acompanhamento do decorrer dos trabalhos, nem sempre facilitado pela distância. A sua visão sobre ciência e contínuos incentivos foram essenciais.
Ao Professor Doutor Henrique Fonseca, co-orientador, estou grata pela disponibilidade e cuidado sempre demonstrados neste período.
À Cristina pela sua amizade e pelo seu trabalho que abriu novas perspectivas no estudo das cardosinas. Para a Sofia e para a Anabela, colegas e companheiras nestes anos, agradeço a amizade e a inter-ajuda sentidas.

Ao Prof. Doutor António Correia por acreditar nas minhas capacidades e pelos desafios lançados entretanto. Ao seu grupo de trabalho, e a todos que por lá passaram, agradeço o acolhimento e a amizade recebidos.
Ao Professor Doutor Euclides Pires, "pai das cardosinas", agradeço a sua ponderação, visão científica e experiência de vida que nos guiou no decorrer deste trabalho.
Ao Professor Doutor Arthur Moir da Universidade de Sheffield pela sua disponibilidade surpreendente e pelo incentivo e optimismo transmitidos.
Ao Professor Doutor Enrique Villar pelo uso do seu laboratório na Universidade de Salamanca, no Departamento de Bioquímica e Biologia Molecular. Ao Professor Doutor Shnyrov agradeço as aprendizagens assimiladas na área da biofísica de proteínas que muito ajudaram na execução deste projecto. Aos colegas David Pina e Anna Shnyrova pela colaboração prestada.
Ao Professor Doutor Francisco Gavillanes por me ter recebido no seu laboratório na Universidade Complutense de Madrid, no Departamento de Bioquímica e Biologia Molecular. Agradeço também a sua disponibilidade bem como os conhecimentos transmitidos sobre dicroismo circular de proteínas.
Ao Professor Doutor João da Costa Pessoa por me ter recebido no seu laboratório no Instituto Superior Técnico de Lisboa, Departamento de Engenharia Química e Biológica, por ter disponibilizado condições e equipamento para realização de ensaios de dicroismo circular.
Ao Professor Doutor Armando da Costa Duarte pela utilização do seu laboratório para estudos de fluorescência e ao Professor Doutor Valdemar Esteves pelo acompanhamento, rigor, interesse e ajudas prestado.
Por último, queria agradecer a toda a minha família pela sua ajuda, apoio e ânimo. Aos meus sobrinhos em especial agradeço a alegria e carinho sentidos na sua companhia. Aos meus pais pela confiança que mostraram ter nas minhas capacidades e em todas as decisões que tenho tomado.
Finalmente, ao João, por tudo.
Este trabalho foi financiado através da bolsa de Doutoramento da Universidade de Aveiro e pela FCT através do projecto POCTI/QUI/60791/2004.

palavras-chave

Cardosina A, Proteases Aspárticas, Estados Conformacionais, pH, Solventes Orgânicos

resumo

As proteinases aspárticas (PAs) são um importante grupo de enzimas devido ao envolvimento em processos patológicos e fisiológicos. Apesar da diversidade funcional apresentam homologia nos seus domínios *N*- e *C*-terminal. Esta estrutura parece ter evoluído através de processos de duplicação e fusão interna de genes. Entre os membros das PAs está a cardosina A de *Cynara cardunculus* sp, uma enzima heterodimérica que tem sido alvo de intensa caracterização. O seu processo de purificação permite grandes rendimentos de enzima pura, tornando-se um candidato atractivo para proteína modelo em estudos de estrutura/função. Nos últimos anos alguns estudos de estrutura de proteínas têm sido feitos sobre os estados desnaturados e não nativos de proteínas. O objectivo destes estudos é a investigação do desenrolamento de proteínas e das possíveis funções fisiológicas destes estados não nativos. As PAs são enzimas derivadas de zimogéneos, que normalmente desenrolam irreversivelmente e cujos estados parcialmente desnaturados são estabilizados. No entanto, ainda há pouca informação sobre os estados não nativos. A caracterização estrutural destas proteínas poderá esclarecer as propriedades funcionais das PAs, bem como a estrutura, o enrolamento e a actividade das proteínas em geral. Neste trabalho os efeitos do pH e do acetonitrilo foram estudados para obter informação acerca dos estados conformacionais da cardosina A. Os estados conformacionais induzidos pelo pH e pelo acetonitrilo mostraram ser diferentes. A desnaturação térmica da cardosina A mostrou tratar-se de uma proteína estável provavelmente devido às fortes interacções entre subunidades que vão enfraquecendo à medida que o valor de pH varia. A desnaturação da cardosina A mostrou ocorrer através do desenrolamento independente das suas cadeias polipeptídicas, onde a cadeia polipéptidica de baixo peso é menos estável do que a de alto peso. Adicionalmente, a desnaturação alcalina da cardosina A resulta na dissociação do heterodímero. Por outro lado, baixas concentrações de acetonitrilo induziram um aumento da actividade da cardosina A relacionada com ligeiro aumento da estrutura secundária. No geral, o acetonitrilo induz alterações de estrutura secundária e terciária na cardosina A inactivando-a a concentrações mais elevadas, mas baixas concentrações do solvente parecem resultar num efeito directo no local activo. Finalmente, as alterações dependentes do tempo em 10 % acetonitrilo mostraram que a cardosina A adopta uma conformação onde o desenrolamento da cadeia de baixo peso é notório. Neste estado a cardosina A apresenta uma actividade catalítica comparável à do estado nativo sugerindo uma elevada flexibilidade funcional e conformacional da enzima. Os resultados foram comparados com informação estrutural e funcional de outras PAs e as possíveis implicações da natureza heterodimérica da cardosina A discutidas.

keywords Cardosin A, Aspartic Proteinases, Conformational States, pH, Organic Solvents

abstract

Aspartic proteinases (APs) are an important group of proteinases due to their involvement in pathological and physiological processes. Despite their function diversity they have *N*-terminal and *C*-terminal domains that show evidence of homology to each other. This structure appears to have arisen by ancient internal gene duplication and fusions. Among the most interesting members of aspartic proteinases is cardosin A from *Cynara cardunculus* sp, a recently well characterized heterodimeric enzyme. High yields of pure protein can be obtained making it an attractive model protein for structure/function studies.
In the last years protein structural studies have been carried out on denatured and other non-native states of proteins. The aim of such studies is to investigate protein folding and to understand the physiological roles of non-native structures. APs are zymogen-derived enzymes, which usually unfold irreversibly and become trapped in partially denatured states. There is still little information about the structures of these denatured states. Characterization of such states in APs should give insights not only into the functional properties of this very important family of enzymes, but also in protein folding and activity of proteins in general.
In this work the effects of varying pH and acetonitrile concentrations on the structure and folding of cardosin A were investigated to obtain further insight of cardosin A conformational states. pH and acetonitrile induced conformational states in cardosin A revealed to be different. It is a stable protein probably due to strong intersubunit interactions that weaken as pH values varies. Furthermore, the unfolding of cardosin A occurs through independent unfolding of its polypeptide chains, where the small chain is less stable than the longer chain. The alkaline denaturation of cardosin A was shown to point to the heterodimer dissociation.
At low acetonitrile concentrations cardosin A activity is increased and correlated with mild secondary structural changes. Overall, acetonitrile induce tertiary and secondary structure changes in cardosin A molecule, causing the complete inactivation at high acetonitrile concentrations and at low organic solvent content a direct effect in the active site was suggested.
Finally, time dependent changes induced by 10 % acetonitrile were studied and showed that at some point cardosin A adopts a partial folded conformation with the unfolding of the small chain. This conformational state is characterized by catalytic activity comparable with the native cardosin A, pointing to high cardosin A functional conformational flexibility. Results were compared with available structure/function information of other APs and the possible physiological roles of the heterodimeric nature of cardosin A discussed.

Parts of this thesis have been published or are in preparation to publish:

Pina, DG*, **Oliveira, CS***, Sarmento, AC, Barros, MT, Pires, E, Zhadan, GG, Villar, E, Gavilanes, F, Shnyrov, VL. 2003. Thermostability of cardosin A from *Cynara cardunculus* L. Thermochimica Acta. 402: 123–134.

* First authors

Shnyrova, AV, **Oliveira, CS**, Sarmento, AC, Barros, MT, Zhadan, GG, Roig, MG, Shnyrov, VL. 2006. Effect of acetonitrile on Cynara cardunculus L. cardosin A stability. International Journal of Biological Macromolecules. 39: 273-279.

Oliveira, CS, Sarmento, AC, Pereira, O, Correia, I, Costa Pessoa, J, Duarte, A, Esteves, V, Fonseca, H, Pires, E, Barros, MT. Activity modulation induced by acetonitrile in cardosin A. Structural and functional perspectives. *In preparation*.

"It is interesting to contemplate a tangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent upon each other in so complex a manner, have all been produced by laws acting around us."

Charles Darwin, The Origin of Species (1859)

Index

4 Results and Discussion

Abbreviations

AEC	Anion exchange chromatography
AP	Aspartic proteinase
APP	Amyloidal protein precursor
ATEE	N-acetil-L-tryptophan ethyl ester
AU	Arbitrary units
Bicine	N,N-bis-(2-hydroxyethyl)glycine
BSA	Bovine serum albumine
C	Heat capacity
°C	Degrees celsius
Cal	Calorimetrically determined value
CD	Circular dichroism
CP	Heat capacity
D	Denatured
DSC	Differential scanning calorimetry
DAN	Diazoacetylnorleucine methyl ester
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
E	Enzyme
EPNE	1,2-Epox-3-(p-nitrophenoxy) propane
FPLC	Fast protein liquid chromatography
G^o	Free energy
H	Enthalpy

HIV	Human immunodeficiency virus
I	Intermediate
K_{cat}	Turnover number (catalytic constant)
kDa	Kilo Dalton
Km	Michaelis-Menten constant
LMW	Low molecular weight
N	Native
RP-HPLC	Reverse phase high performance liquid chromatography
PAGE	Polyacrilamide gel electrophoresis
Pepstatin	Isovaleryl-Val-Val-Statyl-Ala-Statyl
R	Ideal gas constant
Rpm	Rotations per minute
S	Substrate/Entropy
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
Tris	Tris(hydroxymethyl)aminomethane
TEMED	N,N,N′,N′-Tetramethylethylenediamine
TFA	Trifluoroacetic acid
T_m	Midpoint of temperature-induced transition
T_s	Temperature of maximum stability
U	Unfolded
UV	Ultraviolet
V_{max}	Maximum velocity

List of Figures

1 Introduction

3 Material and methods

4 Results and discussion

List of Tables

1 Introduction

1.1. Protein stability

Protein stability concerns the net balance of forces that enables a protein to retain its native conformation under given conditions. Both the strength and specificity of many of these forces closely depend on environmental conditions. Changes in the environment can reduce or eliminate part of the conformational interactions, while the rest are unchanged or intensified and can also induce new conformations with properties intermediate between those of the native and the completely unfolded states (Fersht, 1999).

In what concerns enzymes, protein stability determines and limits their usefulness and this subject has long been of major concern to enzymologists. In fact, the stability must be great enough to maintain the proteins native conformation, but not as great as to preclude conformational changes or adjustments considered essential for many protein functions (Pace, 1990).

1.1.1. Chemical and thermodynamic stability of globular proteins

There are two different perspectives on protein stability, chemical and physical (thermodynamic) stability, normally the term refers to the latter. Chemical stability concerns the protein loss of integrity due to bond cleavage, that is to say due to covalent changes, and is usually an irreversible process. The mechanisms of irreversible protein inactivation often follow common pathways (Volkin *et* a*l*., 1991). The other is the conformational stability of the folded state, in the absence of covalent changes, and is always reversible, with complete recovery of the enzymatic activity, if environment conditions are restored.

The native and catalytically active conformation of an enzyme is maintained by a balance of noncovalent forces (hydrogen bonds, hydrophobic, ionic and van der Waals interactions) (Pace *et* a*l*., 1996), resulting, in a given environment, in the minimum free energy of the entire system (Anfinsen *et* a*l*., 1975). Alterations of this delicate balance by changes of temperature, pH, by salts or by organic solvents addition, may distort

this equilibrium and induce protein molecules to unfold. Usually, quite different changes in the proteins such as loss of activity, changes of spectral properties and other physical parameters or even cleavage of the polypeptide chain can be detected and investigated.

Usually the reversible partial unfolding of a protein is the first step, and if the unfavourable conditions persist, followed by conformational or covalent processes as shown below:

$$N \xleftarrow{K_U} U \xrightarrow{K_I} I \qquad 1$$

where N is the native form of the protein undergoing unfolding, U corresponds to the reversible unfolded form and I relates to the irreversibly inactive protein fraction.

Several chemical reactions leading to irreversible enzyme inactivation have already been described (Volkin *et al.*, 1991). For example at low pH hydrolysis of peptide bonds at Asp residues (the main process) and deamidation of Asn and/or Gln residues are favoured. At neutral pH inactivation is caused by a combination of disulfide interchange (the main process), beta-elimination of Cys residues, and deamidation of Asp and/or Gln residues. Other processes include thiol-catalyzed disulfide interchange and oxidation of Cys residues.

To quantify protein stability the conformational stability is commonly used and, therefore, described on thermodynamic quantities. The thermodynamic stability of a globular protein can be defined as the difference in the free energy between the folded and unfolded conformations under standard conditions and denoted as $\Delta G(H_2O)$. This stands for the Gibbs energy change for the equilibrium reaction taking place without cleavage of covalent bonds. Therefore it is possible to compare proteins independently of their biological activity, molecular mass, etc. Moreover, related thermodynamic quantities such as enthalpy change (ΔH), entropy change (ΔS), and heat capacity change (ΔC_p) at unfolding can be determined which allows gaining a deeper insight into the forces that stabilize the unique native three-dimensional structure (Privalov, 1979).

1.1.2. From monomeric to multimeric proteins

1.1.2.1. Monomeric proteins

The studies on protein stability initially focused on small reversibly unfolding globular proteins, mostly with disulfide bonds intact (Pace, 1990; Privalov, 1979) and with their three dimensional structure solved. Equilibrium denaturation and kinetic studies have enabled analysis of packing forces in proteins, testing the globular folding of mutant proteins as well as the functional interactions of residues further stimulating enzymology as well as biotechnology (Dill *et* a*l*., 1995; Mathews, 1993). It has been calculated that the range of stabilities of most naturally occurring small monomeric globular proteins (from ~ 50 to 200 amino acids) lies between 5 and 15 Kcal/mol and represents the small difference between multiple noncovalent interactions favouring the folded protein structure and unfavourable entropic terms (Pace, 1990). At equilibrium unfolded transitions of single domain proteins are usually two-state as shown below:

$$N \xleftrightarrow{K} U \qquad 2$$

with only the fully folded (*N*) and unfolded (*U*) states populated, even though a subset of *U* states can exist (Privalov, 1979). Other conformations such as partially-folded conformations with distinct thermodynamic properties are energetically unstable. At this point there is a single rate constant (*k*), so all the molecules unfold with the same probability. More studies were subsequently conducted with larger proteins with multi-domains that have been shown to unfold step-wise, with the domains unfolding individually (Privalov, 1982), either independently or with varying degrees of interactions between them (Brandts *et* a*l*., 1989; Griko *et* a*l*., 1989) and often presenting stable intermediates. These partially folded sates in monomeric globular proteins have been correlated strictly with the bulk content of hydrophobic and charged residues and not with its amino acid sequence. In this study, proteins that do not have equilibrium

intermediates were shown to be less hydrophobic and to have, in general, a larger net charge than those capable to form intermediate sates (Uversky, 2002). Other investigations also reported the importance of internal empty and/or water filled cavities (Williams *et* a*l*., 1994).

At this point it is clear that proteins in general can adopt different stable partially folded conformations. These have been considered to play crucial roles in protein synthesis and folding, function and degradation of globular proteins (Tcherkasskaya *et* a*l*., 1999; Uversky, 1997). Moreover, the aggregation of partially folded proteins is associated with a number of human diseases (Ramos *et al*, 2005; Uversky *et* a*l*., 1998), and often constitutes a significant problem in biotechnology (Singh *et* a*l*., 2005).

1.1.2.2. Oligomeric proteins

All these stability studies with monomeric proteins and also the development of biophysical and spectroscopic techniques to follow structural changes in proteins encouraged stability studies of oligomeric proteins. They presented a challenge and an excellent opportunity to learn about the importance of quaternary structures together with crystallographic data that also provided important information on the conformational stabilization of the subunits interface. In fact, with dimeric or oligomeric proteins additional modes of stabilization (intrachain and interchain interactions) have seen to be available at the quaternary structural level (Jones *et* a*l*., 1995; Jones *et* a*l*., 2000).

The unfolding pathways of oligomers are often complicated (Neet *et* a*l*., 1994). Many dimeric proteins denature with a two-state equilibrium transition, whereas others have stable intermediates in the process. For those proteins showing a single transition of native dimer to unfolded monomer (two-state), the conformational stabilities range from 10 to 27 Kcal/mol, which is significantly greater than the conformational stability found for monomeric proteins, suggesting that the stabilization energy of dimers is primarily due to intersubunit interactions (Brandts *et* a*l*., 1989). Concerning the transitions with stable intermediates, these multisubunit proteins usually dissociate first, with subsequent subunits unfolding, unless specific conditions are present, like

domain peripheral location in the molecule, promoting the independent unfolding of the domains (Jaenicke, 1987). At equilibrium unfolded transitions of oligomeric proteins can be represented by the general 3-state equation as shown below:

$$N_2 \xleftrightarrow{k_1} 2N \xleftrightarrow{k_2} {}_2U \qquad or \qquad N_2 \xleftrightarrow{k_1} I_2 \xleftrightarrow{k_2} {}_2U \qquad 3$$

where N_2 is the native dimer, N the native monomer or a monomeric intermediate (I), U the unfolded monomer and k representing equilibrium constants. With the exceptions of non-specific aggregation of unfolded intermediates or covalently linked dimers, completely unfolded dimers are not likely to occur. Nevertheless, partially unfolded dimeric intermediates (I_2) have been reported (Brandts *et* a*l*., 1989, Blackburn *et* a*l*., 1981). Finally, and although dimeric proteins have additional modes of quaternary structure stabilization, many still follow a 2-state transition, represented in the equation above (Equation 3) when dissociation leads to the formation of an intrinsically unstable species and consequently a folded monomeric species will not be significantly populated at equilibrium.

1.1.2.2.1. The evolution and biological value of oligomeric proteins

Currently, the Protein Data Bank (PDB) contains more than 2000 oligomeric structures out of around 40 000 structures available. Oligomers represent an important portion of the existing proteins and are capable of performing tasks for which monomers are not suited for. They often work as the functional form of proteins and therefore present biological advantages. The existence of multienzymes complexes with their assembly line production efficiency, or enzymes that can produce a catalytic site only upon subunit association, and enzymes that have evolved in different organisms into different subunit compositions displaying different substrate specificities, are also examples of the flexible response to diverse environments assigned to oligomers. Finally, dynamic rather than stable, protein oligomerization has also been described for biological functions such as the regulation of DNA expression, or the transduction of

signals across the cell membrane (promoted either by protein association or dissociation) and also to perform the catalytic cycle of ATP synthesis in mitochondria depending in the continuous oscillation between alternative oligomeric states (D'Alessio, 1999).

The advantages of oligomeric proteins compared to their non-associated counterparts, however, can also be conferred by multi-domain proteins. There seems to be no fundamental structural and functional distinction between domains and subunits. On the other hand, and in the case of large multidomain proteins, their occurrence is rare, being most large protein structures formed by association of subunits. Oligomerization of small gene products is by far more economical, as observed in viruses where a small gene is sufficient for building up a large virus capsid. Furthermore, errors occurring during protein synthesis can be eliminated more easily in the case of oligomeric structures.

Since the awareness of the biological importance of oligomers, efforts have been made to understand the mechanism of oligomerization and of the evolution of protein oligomers. This was facilitated by the increasing availability of suitable case studies. In fact the evolutionary transition can only be studied by investigating genetics and structure of pairs of homologous proteins, usually, each composed of a monomeric and a dimeric counterpart. Some protein pairs have become model systems in these studies such as the monomeric bovine pancreatic RNase and the evolutionary related dimeric seminal RNase A (D'Alessio, 1990b) and also the homologous monomeric γ–type and oligomeric β–type crystallins (D'Alessio, 2002).

The hypothesis that oligomeric proteins emerged first as functional aggregates and later dissociated into functional monomers is acknowledged but additionally there are evidences that divergent evolution more often used the association of protein units into oligomers to vary and enrich the cell repertoire of structures and functions. This can be clearly seen in the "hydrophilic effect" recorded at intersubunit interfaces, where a significant presence of polar and charged residues at oligomeric interfaces is seen. This can result from the association of previously exposed, hydrophilic surfaces (from a monomer) into solvent-excluded interfaces (in an oligomer) (D'Alessio, 1999). In general both structural and mutation studies suggest that the association of monomers into oligomers could result from one or more mutational events in the monomeric ancestor followed by swapping of domains where interdomain interfaces would be

readily reconstituted as intersubunit interfaces (Jones *et* al., 2000). Whatever the exact mechanism occurring in the evolutionary processes for oligomerization, it seems reasonable to assume that there is a step zero in which expendable genetic material becomes available to the genome (either by gene duplication or some other genetic event). In this way no loss of structure and function occurs in the organism in which the oligomerization takes place. Furthermore it is expected that the new structure obeys thermodynamics and chemistry rules, be stable and maintained by appropriate chemical bonds and forces, and that it is biologically significant to the organism in which the evolutionary event takes place.

Besides investigations of mutational events by amino-acid substitutions in homologous proteins, another tool is useful in shedding light on putative ancestors of present-day protein oligomers. It has been concluded that the analysis of the refolding mechanism by which denatured, unfolded polypeptide chains fold back into oligomers may shed light on the evolutionary history of the oligomers, as this might be recapitulated in the pathway of oligomer refolding (D'Alessio, 1999).

Although the studies on proteins of similar structures and homologies have not yielded definite answers concerning the subject of the evolutionary transition to oligomeric, more of these studies should be undertaken. The underlying issue of solving the protein folding problem (Uversky, 2002) is always present. At present, only few studies have been performed to produce a definitive conclusion or to formulate theories to explain the diverse folding behaviours that have been observed in family comparisons. As more comparative studies of proteins with similar structures are completed, the types of protein structures, or even substructures within various proteins, that tend to display conserved folding mechanisms will become apparent (Zarrine-Afsar *et* a*l*., 2005). These data may ultimately allow us to fully understand the influences of topology, tertiary packing and structural propensities on the folding pathways of all proteins. In fact, trying to fully understand the ambiguous answers in comparative folding studies could ultimately provide the deepest insight of the folding code as well as of protein evolution issues.

1.2. Aspartic proteinases

Aspartic proteinases (APs) (EC.3.4.23) constitute one of the four major catalytic types of proteases. In MEROPS database (http://www.merops.ac.uk) (Rawlings *et al.*, 2006) there are, at present, and from a total number of 48970 registered protease sequences with 383 PDB entries, 2915 registered APs sequences, from which 36 have their three dimensional structure solved. APs are then grouped into 14 different families, a set of homologous proteins, regarding significant similarity in amino acid sequences (Table 1). Families are also divided into subfamilies when there is evidence of ancient divergence within the family by divergent evolution. At the same time, families can be grouped in clans (7 in AP, Table 1.1) when there is evidence of single evolutionary relationship. These can be the similarity of tertiary structures, or when structures are not available, by the order of catalytic-site residues in the polypeptide chain and often by common sequence motifs around the catalytic residues.

Despite the sequence homology APs are widely distributed in a variety of organisms like in virures, bacteria, archaea, archezoa, protozoa, plants and animals. However little is still known about the majority of the families, being family 1 and 2 the most studied, both belonging to clan AA.

Table 1.1: Overview of APs families (Rawlings *et al*., 2006).

Aspartic proteases (EC.3.4.23)

Family /Clan	Type example	Content of family	Distribution	Molecular organization	Biological functions
A1/ AA	Pepsin A (Homo sapiens)	Endopeptidases, most of which are most active at acidic pH.	Protozoa Fungi Plants Animals	Monomeric and oligomeric (homodimeric and heterodimeric)	Synthesised with signal peptides, and the proenzymes are secreted into the lysosomal/endosomal system, where acidification leads to autocatalytic activation.
A2/ AA	HIV-retropepsin (human immunodeficiency virus 1)	Endopeptidases.	Fungi Animals Virus	Homodimeric	Retropepsin is released from the *gag-pol* polyprotein autolytically and the peptidase has to dimerize to be active. It is thought to follow release of retropepsin monomers from the viral polyprotein by host peptidases.
A3/ AA	Cauliflower mosaic virus-type peptidase (cauliflower mosaic virus)	Endopeptidases	Plants Virus	No tertiary structure has been determined. Assumed to be a homodimer	Polyprotein-processing endopeptidases of the pararetroviruses, double-stranded DNA viruses that infect plants.
A5/ A-	Thermopsin (Sulfolobus acidocaldarius)	Endopeptidases	Archaea (apparently restricted)	No tertiary structure has been determined.	*Sulfolobus acidocaldarius* is a thermophilic archaean. Thermopsin is probably important for the nutrition of the organism.
A6/ AB	Nodavirus peptidase (flock house virus)	Endopeptidase	Virus (known only from the genus *Nodavirus)*	Two different domains	The coat protein precursor undergoes autolytic maturation stabilizizing the virion.
	Signal peptidase II	Endopeptidase	Bacteria	No tertiary structure has	Signal peptidase II is essential for the maturation of the

A8/ AC	(*Escherichia coli*)		Animals	been determined.	premurein lipoprotein in the bacterial cell wall.
A9/ AA	Spumapepsin (human spumaretrovirus)	Endopeptidase	Virus (known only from the genus *Spumaviru*)	No tertiary structure has been determined.	Processing of the Gag polyprotein by spumapepsin is essential for viral infectivity: without it viral cDNA is not synthesized.
A11/ AA	Copia transposon (Drosophila melanogaster)	Endopeptidases	Fungi Plants Animals	Fold is assumed to be similar to that of retropepsin.	The LTR-retrotransposon encodes an endopeptidase containing polyprotein that releases the proteins from the polyprotein. Some retrotransposon endopeptidases are similar to retropepsin and included in the same family.
A21/ AB	Tetravirus peptidase (Nudaurelia capensis omega virus)	Tetravirus endopeptidases	Virus	The fold is similar to that of the known structures from family A6.	Tetraviruses are single-stranded RNA viruses that infect insects.
A22/ AD	Presenilin 1 (*Homo sapiens*)	Membrane inserted endopeptidases	Archaea Protozoa Fungi Plants Animals	Eight predicted transmembrane domains even though protein fold and active site residues are not known.	Following the processing of the beta-amyloid precursor protein by beta-secretase, the presenilins as part of gamma-secretase cleave the membrane bound C-terminal of the precursor protein within its transmembrane region.
A24/ AD	Type 4 prepilin peptidase 1	Membrane-inserted endopeptidases	Bacteria Archaea	Eight predicted transmembrane domains	Type IV pilus formation, toxin and other enzyme secretion, gene transfer and biofilm formation.
A25/ AE	Gpr peptidase (Bacillus megaterium)	Resemblance with HybD endopeptidase	Bacteria (*Firmicutes* only)	-	-
A26/ AF	Omptin (Escherichia coli)	Membrane-inserted endopeptidases	Bacteria (only in proteobacteria)	Ten-stranded anti-parallel beta-barrel structure.	May have a role in an adaptive immune response, enabling survival *in vivo*.
A31/ AE	HybD peptidase (*Escherichia coli*)		Bacteria Archaea	-	-

Tertiary structures have been determined for members of family A1, divided into 2 subfamilies (A1A and A1B). Up until 20 years ago the known sources of aspartic proteinases were stomach (pepsin, gastricsin and chymosin), lysosomes (for cathepsin D), kidney (for renin), yeast granules, and fungi (for secreted proteases like rhizopuspepsin, penicillopepsin and endothiapepsin). By then four aspartic proteases crystal structures were solved at high resolution: porcine pepsin (Northrop, 1930), the first enzyme to be discovered and the first crystal structure of an aspartic to be solved, rhizopuspepsin (Ohtsuru *et* a*l*., 1982), penicillopepsin (Emi *et* a*l*., 1976) and endothiapepsin (Whitaker, 1970), all grouped in sub-family A1A (Table 1.2). In fact, there are still no PDB entries of family A1B proteases sequences, being the subfamily type represented by nepenthesin from carnivorous plant and believed to be responsible for digestion of insect prey in the genus *Nepenthes* (Kenji *et al*., 2005). In family A2 (Table 1.1), where retroviral APs are grouped (Seelmeier *et* a*l*., 1988), dimerization has to occur to form an active peptidase, producing a structure very similar to that of pepsin (family A1), except that the single flap (explained below) is absent.

Table 1.2: List of proteinases and homologues from sub-family A1A from APs, clan AA with PDB entries (Rawlings *et al*., 2006).

Sub family	Proteinase[a]	Source	State	Localization	Function	PDB[b]
A1A	Pepsin-A	Mammalian	Monomeric	Lumen stomach	Proteolysis in the stomach.	5PEP; 1PSN; 1FLH
	Gastricsin					1HTR[c]
	Memapsin-2			Transmembrane (endossomes)	Primary beta-secretase activity in mouse brain.	1W50
	Chymosin			Lumen stomach	Digestion of *k*-casein in neonatal gastric digestion.	4CMS
	Renin				Catalyzes the first step in angiotensin II.	1BBS
	Renin-2			Endothelial and lumen vessels.	production.	1SMR[d]
	Mername-AA047 peptidase			Stomach mucosa	Proteolysis in the stomach.	1AM5
	Cathepsin D		Heterodimeric	Lysossomes	Contributes to lysosomal proteolysis.	1LYA
	Cathepsin E			Lysossomes	Processing of protein antigens	1LCG[e]
	Penicillopepsin	Fungal		Extracelular	Excreted into the surrounding environment where hydrolyses proteins.	3APP
	Rhizopuspepsin					2APR
	Mucorpepsin					1MPP; 2ASI
	Aspergillopepsin					1IBQ
	Endothiapepsin					4APE
	Saccharopepsin					1FMU
	Polyporopepsin					1WKR
	Candidapepsin					
	SAP2					1ZAP[d]
	oryzepsin					1IZD
	Canditropsin					1J71
	Yapsin-1			Plasma membrane or the cell wall	Participation in cell-wall assembly and/or remodelling	1YPS[e]
	Phytepsin	Plant	Heterodimeric	Storage vacuoles	Protein processing/degradation in different stages of plant development	1B5F
	Plasmepsin	Protozoa	Monomeric	Digestive vacuoles	Important in haemoglobin degradation.	1LF4
	Plasmepsin-1					1LCR[e]
	Plasmepsin-2					1PFZ
	Plasmepsin-4					1LS5

a) Only proteinases and homologues with PDB entries.
b) Unless noted, PDB code refers to three dimensional data of mature wild type protein forms.
c) Precursor form.
d) Complexed with inhibitor form.
e) Theoretical model.

1.2.1. General characteristics

With the increasing number of aspartic proteinases being characterised from vertebrates, insects, helminths, protozoans, plants, retroviruses and bacteria it is becoming clear that further than amino acid sequence homology between AP, not always tertiary structure homology is observed, as can be seen in Table 1.1. Thus it is not easy to produce a set of rules or characteristics that apply to all of these enzymes. In general pepsin, the clan AA type proteinase, is considered to be the archetypal aspartic proteinase and therefore only pepsin-like APs, will be discussed here.

The typical characteristics of these enzymes are pH optima for catalytic function in the acid range and their inhibition by pepstatin (a hexapeptide from *Streptomyces*) with the exception of the pepstatin insensitive carboxyl proteinases (Oyama *et* a*l*., 1999). They also present strong preference to cleave peptide bonds between large and hydrophobic residues, with the exception for yapsins, a yeast proteinase from *Saccharomyces cerevisiae*, from sub-family A1A APs having shown preference for basic residues (Ledgerwood *et al*, 1996; Gagnon-Arsenault *et al*., 2006).

Peptidases in family A1 are all-beta proteins consisting of two similar beta barrel domains, each of which contributes one active site Asp, as can be seen in Figure 1. These catalytic aspartates are positioned within the highly conserved Asp32-Thr-Gly or Asp215-Ser-Gly motifs (DTG and DSG respectively) on which grounds APs classification (Tang *et* a*l*., 1987). Besides this identity, the vicinity is also typical consisting in more detail in *Xaa-Xaa-Asp-Xbb-Gly-Xbb*, where *Xaa* represents a hydrophobic residue and *Xbb* represents a residue of either serine or threonine (Barrett *et* a*l*., 1998). Concerning highly conserved amino acid residues, Tyr75 and Thr77 (pepsin numbering) are also frequent and located in a very mobile loop (flap) projecting out over the active site cleft and interacting with the protein substrate (Figure 1.1) (Davies, 1990). This flap, as mentioned above, is absent from retroviral proteinases (Rao *et al*., 1991). Most of the eukaryotic APs are about 330 amino acids long, whereas retroviral counterparts are considerably smaller having less than 130 amino acids.

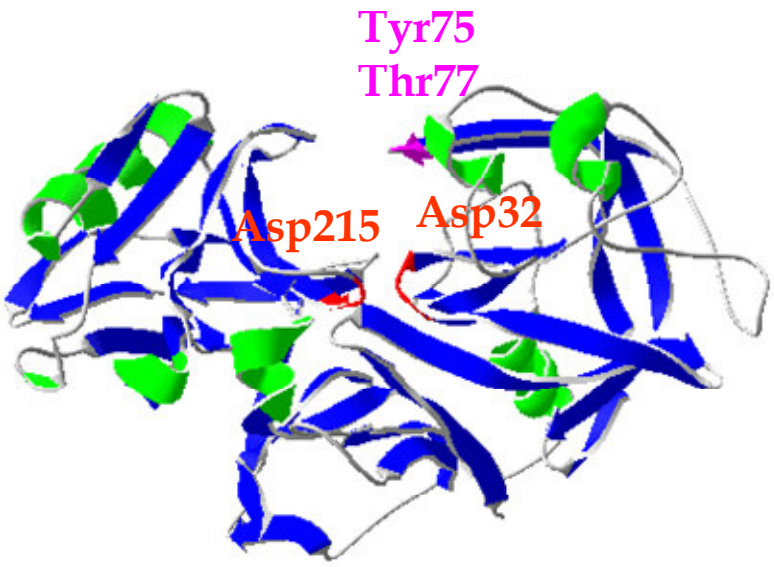

Figure 1.1: Molecule representation of an archetypal aspartic proteinase (pepsin A, pdb code 5pep). Secondary structure elements are represented in blue (strands) and green (helices). Catalytic aspartates, Asp215 and Asp32, are represented in red and Tyr75 and Thr77 in pink.

1.2.2. Synthesis and mechanism of activation of aspartic proteinases

Enzymes that "digest" proteins, like APs, cause a real challenge. The enzyme must be constructed inside the cell, but its activity controlled so that it doesn't immediately start digesting the cell's own proteins. To solve this, pepsin and many other protein-cutting enzymes are created as inactive "proenzymes," which may then be activated once safely outside the cell. In this sense, all non viral APs are synthesized as inactive precursors (zymogens). Zymogen conversion to the active enzyme generally occurs by limited proteolysis of an inhibitory activation segment within a subcellular compartment or the extracellular environment, in which the particular enzyme functions. Conversion may involve accessory molecules, or the process to be autocatalytic, requiring a reduction in pH. In APs of non viral origin the zymogens

have *N*-terminal extensions (propeptide) bound to the active site cleft, preventing undesirable degradation during intracellular transport and secretion (Dunn, 2002; Khan *et* a*l*., 1998). Furthermore it has been suggested that they also play a role in the correct folding, stability and intracellular sorting (Koelsch *et* a*l*., 1994). The *N*-terminal propetide is usually about 50 amino acids long even though plasmepsins have shown to have propeptides with about 120 amino acids. The propeptide release, which activates the zymogen, is carried out at low pH resulting in proteolytic cleavage of a pro-part (Rawlings *et* a*l*., 1995). Several three dimensional structures of zymogens have been reported, globally being detected distinct mechanisms of zymogens inactivation. Frequently it is achieved by blocking the preformed active site, by the positively charged residues of the propeptide, or by the Lys/Tyr residues from the mature enzyme. The first relates to pepsinogen, progastricsin, prochymosin and procathepsin *D* forms (Ritcher *et* a*l*., 1998; Wittlin *et* a*l*., 1999), and the latter to prophytepsin from barley (Kervinen *et* a*l*., 1999). A third process and more unusually observed occurs in proplasmepsin where the propeptide intertwines around the *C*-terminal domain, causing a distortion in the active site, precluding the general base activation of the nucleophilicic water molecule.

Additionally, the cDNA cloning of some plant APs has demonstrated the presence in the *C*-terminal region of an insert of approximately 100 non homologous amino acids not found in animal or microbial APs. Indeed they are synthesized as the preproform. This specific insert characterizes APs of plant origin, being known as the plant specific insert (PSI). As to the processing of plant APs, detailed reports have been presented (as an example: Kervinen *et* a*l*., 1999; Faro *et* a*l*., 1990; Domingos *et* a*l*., 2000). According to these reports, processing of plant AP precursors leads ultimately to the formation of a two-chain enzyme, without the prosegment and the PSI domain (Simões *et* a*l*., 2004).

On the other hand retroviral APs, like all the proteinases from viral origin, are synthesised as a part of a structural polyprotein and flanked by structural and enzymatic modules that are required for replication of the virus. The proteolytic enzyme is then excised from the polyprotein by a stepwise and coordinated activation mechanism, still not fully described and understood (Krausslich *et* a*l*., 1988).

1.2.3. Three dimensional structure and the catalytic apparatus

High resolution X-ray structures of APs of many sources have been studied, from retroviruses, protozoa, fungi, plants, fish and mammals, and structures thus obtained. As can be seen in Figure 1.1, the molecule is bilobal consisting of two domains of similar structure related by an approximate dyad rotation axis. Within the *N*- and *C*-terminal domains, less precise two fold symmetry is observed (Whitaker, 1990). Each domain is dominated by packed sheets with several small helical elements. Beneath the two domains, an extensive six-strand antiparallel pleated sheet forms the base of the molecule, each domain contributing with three strands (Davies, 1990).

Observation of the several three dimensional structures available showed that pepsin like APs have different oligomeric states (Table 1.2). There are monomeric (like pepsin) and heterodimeric states (like cathepsin D or phytepsin) as a direct result of the mechanism of zymogen activation explained above and, also, there are homodimeric APs (like in retroviral proteases, in family A2). As can be seen in Figure 1.2, where some APs three dimensional structure are represented, the heterodimeric nature of some APs does not hamper the typical pepsin-like fold, with the chains intertwining similarly.

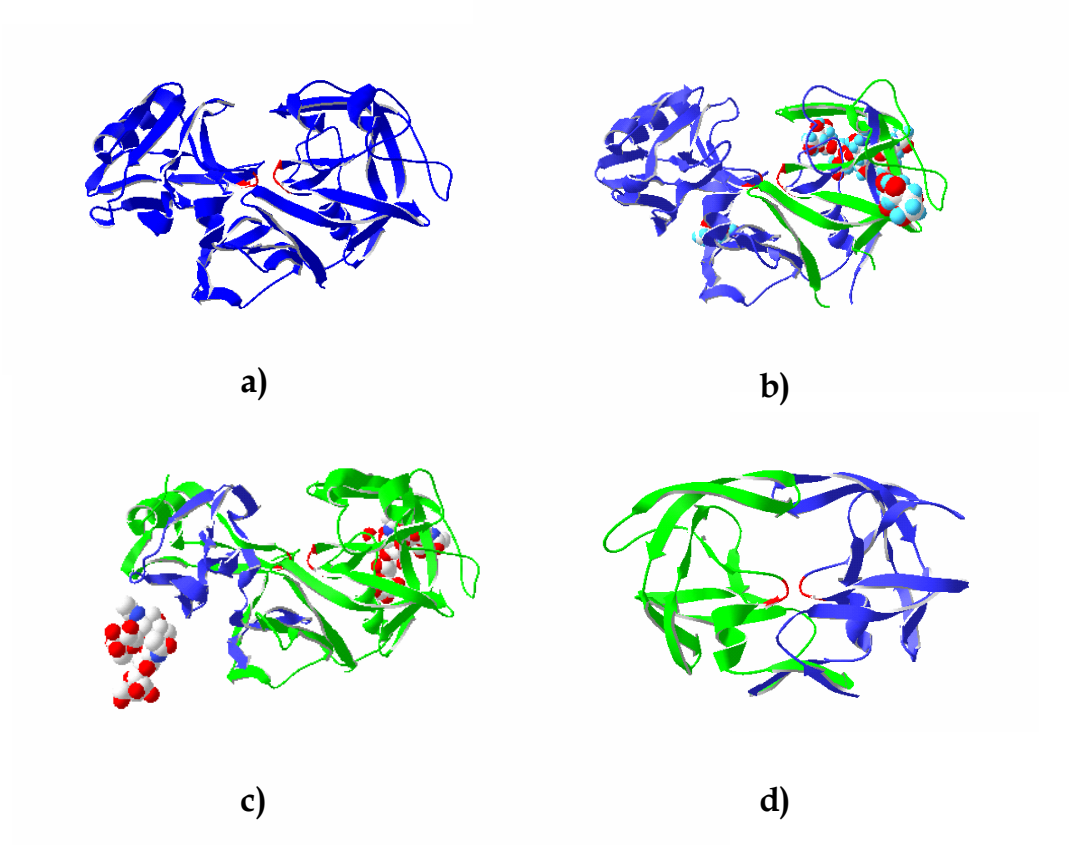

Figure 1.2: Crystal structures of unbound APs. a) porcine pepsin (pdb code 5pep); b) cathepsin D (pdb code 1lya); c) cardosin A (pdb code 1b5f) and d) HIV-1 (pdb code 10dw). Polypeptide chains are represented in blue and green. The catalytic aspartates are represented in red and glycosylation is shown in space-filling form.

The extended cleft (about 40 Å long) can accommodate at least eight residues of a substrate or of an inhibitor in the *S4-S3'* sub-sites having a strong preference for peptide bond cleavage between hydrophobic residues occupying the *S1-S1'* sub-sites (Dunn *et al.*, 200).

The same structure of the catalytic site and the surrounding region in all pepsin like enzymes implies a universal mechanism of their action represented in Figure 1.3. The side-chain carboxyl groups of the aspartates are held coplanar and within hydrogen bonding distance. A solvent molecule is assumed to take part in the catalytic mechanism (Pearl *et al.*, 1984) as represented in Figure 1.3.

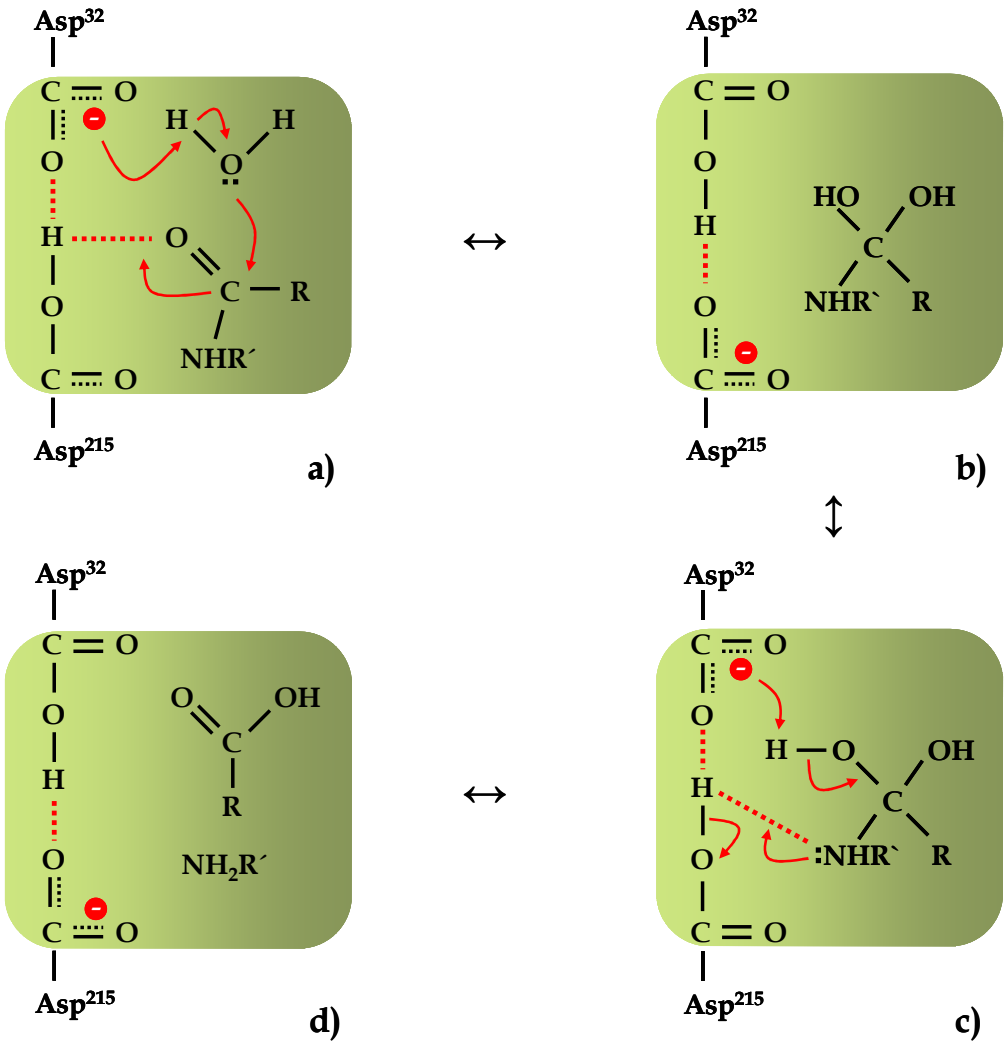

Figure 1.3: Representation of the mechanism of aspartic acid protease catalysed peptide cleavage.

In accordance with the accepted mechanism of pepsin-like enzyme function, Asp215 has to be charged, whereas Asp32 has to be protonated. Regions adjacent to the catalytic centre preserve the charged and protonated states. The water molecule bounded between the active carboxyls, becomes deprotonated on substrate binding to initiate the general base catalysis (Suguna *et al.*, 1987). In addition, another water molecule has found to be completely conserved in pepsin-like enzymes being essential in the formation of a chain of hydrogen-bonded residues between the active-site flap

and the active carboxyls on ligand binding (Andreeva *et al*., 2001). This intricate network renders the active site area rigid as shown by the low temperature *B*-factors observed (Berman *et al*., 2000). On the other hand the flap is characterised by high mobility, as indicated by high *B*-factors. This mobility is needed for closing down over inhibitors bound in the cleft, further excluding solvent, while becoming considerably less mobile (Kempner, 1993; Okoniewska *et al*., 1999).

Therefore, the proposed mechanism of action of the pepsin-like APs involves a first step of protonation of the carbonyl oxygen of the substrate by the proton shared by the two aspartates. The nucleophilic attack on the carbonyl carbon by water with the transfer of a proton to Asp32 follows (Figure 1.3a), resulting in the formation of the tetrahedral intermediate (Figure 1.3b). Products are then formed through the protonation of the nitrogen atom, either from a solvent molecule or from the catalytic carboxyls (Figure 1.3, c and d) (Davies, 1990; Andreeva *et al*., 2001; Veerapandian *et al*., 1992).

1.2.4. Pharmaceutical and biotechnological relevance

In the past years the understanding of the structure, function and diversity of the APs increased greatly. Initially the knowledge gained came mostly from family A1 with the individual enzymes performing many functions and having revealed both biological relevance and pharmaceutical interest. Some like chymosin, rhizopuspepsin, phytepsin and cardosin A have economical importance in the food industry as of their use to coagulate casein in cheese-making and in soya and cocoa processing (Drohse *et al*., 1989; Faro *et al*., 1992).

The most known AP, pepsin, is found in gastric tract of mammals along with secreted acid and it is related with peptic ulcer disease (Cooper, 2002). Furthermore, pepsin has been studied concerning its potential to catalyze peptide synthesis and was shown to be particularly efficient to catalyze peptide bond formation according to equilibrium controlled procedure (Bemquerer *et al*., 1994). Synthesis of peptides is a very important matter in several industries such as for pharmacy and food industry.

Another pepsin like AP of special importance is renin acting in the control of blood pressure and sodium. Although it remains an excellent potential drug target, the problems of poor oral bioavailability and rapid excretion of peptide drugs have precluded effective results. Nevertheless, many of the inhibitors developed proved to be successful lead compounds in the search for inhibitors of HIV proteinase (Cody, 1994; Cooper, 2002).

Cathepsin *D* is a lysosomal AP with the fundamental role of degrading intracellular and internalized proteins. Unlike other members of APs, which are mostly secretory proteins, procathepsin *D* is sorted to the acidic environment of lysosomes where it is activated to a mature two-chain cathepsin *D*. It has been associated in antigen processing and in enzymatic generation of peptide hormones. In the mammary gland the function of cathepsin *D* seems to be related to the processing of the hormone prolactin. Furthermore, increased levels of cathepsin *D* (both at the mRNA and protein levels) were correlated with several human neoplasic tissues and indicated as a growth factor. Therefore, the design of antagonists might be a valuable tool in breast cancer inhibition (Rochefort, 1990). At the same time the observation of high levels of enzymatically competent lysosomal proteases (like cathepsin *D*) abnormally localized in senile plaques in brains of Alzheimer's disease patients represented evidence for candidate enzymes that may mediate the proteolytic formation of amyloid. Consequently modulating the activity of lysosomal proteases may have therapeutic value in Alzheimer disease (Cataldo *et* a*l*., 1990; Cooper, 2002).

Alzheimer's disease is one of a wide range of so-called amyloid diseases which are caused by the formation of insoluble deposits of aberrantly folded proteins in the brain. This protein has been identified as the amyloid *β*–protein (*Aβ*) and results from amyloidogenic processing of a *β*–amyloid precursor protein (APP). The identification of a membrane-bound enzyme (memapsin 2 = BACE) with the ability to cleave the APP provided major impetus for studying the causative factors in Alzheimer's disease. At the same time the designing of drugs which could inhibit this enzyme represents a potential therapy for Alzheimer's disease (Yan *et* a*l*., 1999; Hong *et* a*l*., 2000; Durham *et al*., 2006).

The genome of *Candida albicans* contains ten distinct genes encoding extracellular APs (SAPs = secreted APs) and at least one encoding an intracellular one. It has been demonstrated that the secretion of APs by *Candida* constitutes one of its virulence

determinants encouraging the development of drug targets for *Candida*'s infections treatments (Pichova *et* a*l*., 2001).

Up until now two APs, plasmepsins I and II, from the most virulent human malaria parasite *Plasmodium falciparum* have been identified. These AP and other cysteine proteinases seem to mediate processes in the parasite's life cycle, including the invasion and rupture of erythrocytes and the degradation of haemoglobin by trophozoites. Inhibitors of both AP and cysteine proteinases are under study as potential candidates for anti-malarials and have already showed in vitro to block parasite development and cured malaria-infected mice (Silva *et* a*l*., 1996; Cooper, 2002).

Besides all the abovementioned biological value of family A1 of AP, retroviral AP have also been intensively studied because of its importance in the maturation and infectivity of the human immunodeficiency virus (HIV). It has been observed that mutant viruses containing catalytically inactive proteinase fail to mature and are not infectious predicting successful HIV therapies with inhibitors of retropepsin. In fact, this enzyme is one of the most studied peptidases, with more than 500 sequence recorded and 6 PDB entries (Rawlings *et* a*l*., 2006) and has been the subject of intense research in the development of inhibitors of retropepsins as antiretroviral agents (Wlodawer *et* a*l*., 1998; Dash *et al*., 2003).

There is no question concerning the involvement of APs in important biological processes and the exploration of their potential for biotechnological applications. Their structure homology and the diversity of physiological roles make them interesting model enzymes for structural/function studies. As a result it would not be surprising that new applications will continue to be discovered and old ones improved.

1.2.5. Evolution and adaptation of AP family

From the available AP structures identical folding patterns are observed albeit the diversity in their physiological roles (Tables 1.1 and 1.2). The inhibition by pepstatin, diazoacetylnorleucine methyl ester (DAN) and 1,2-epox-3-(*p*-nitrophenoxy) propane (EPNE) and by covalently reacting inhibitors support this conformational similarity

(Tang, 1971; Rajagopalan *et al*., 1966). On the other hand this functional diversity in APs is believed to be caused by subtle differences of the catalytic apparatus. Here, the intricate network of hydrogen bonding allows physiological adaptations of the same structural motif for the action in diverse conditions representing clearly a gain in evolution. Indeed, a remarkable property of this catalytic center is the adaptation for the action in a wide range of pH from pH 1 up to 7 (Davies, 1990; Dunn, 2002). On contrary, simple homodimeric structures of retroviral aspartic proteinases do not display such a regulating system being their function in cells limited to a narrow interval of pH (pH± 3.5 and pH± 6.5) common to all of them and not far from the pK_a of the carboxyl groups (Hyland *et* a*l*., 1991; Ido *et al*., 1991).

The adaptation of the catalytic apparatus and the folding pattern, namely the two-fold symmetry, indicate that gene duplication and fusion events have resulted in eukaryotic APs that evolved divergently from ancient dimeric enzymes, resembling retroviral APs. Up until now no clear primordial enzyme has yet been observed (Tang *et* a*l*., 1987).

The chase for indications on the selective forces responsible for the evolutionary process that occurred in the pepsin gene family has already been attempted (Carginale *et al*., 2004). Analysis of the gene phylogeny built upon nucleotide sequences coding for various pepsins in 30 vertebrate species showed that positive selection occurred after gene duplication leading to the formation of distinct pepsin groups. This is frequently observed in other gene families evolution investigations (Bielawski *et al*., 2000; Seoighe *et al*., 2003). Also the occurrence in the pepsin family of adaptative evolution in the absence of recent duplication events was suggested and may reflect the existence of long-term selective pressure. In conclusion, a complex evolutionary pattern of the pepsin gene family was found, characterised by some gene duplication events and possibly also by gene loss. Usually, duplications are followed by positive selection, but in some cases the latter may have occurred also long after divergence. At the protein level, gene duplication is characterized by functional divergence involving modifications of specific amino acid sites.

Concerning the mono or heterodimeric nature of APs, apparently in family A1 (Table 1.2) the distribution of monomeric and heterodimeric proteases does not follow a clear trend, having in mind their source, localization and function, at least with the data available so far. In Table 1.2 only APs with solved three-dimensional structure by X-ray

crystallography are represented, summing 25 proteins with PDB entries in the subfamily A1A, from out of 731 sequences known to date (Rawlings *et al*., 2006). Getting information regarding the structure of the remaining proteases will help solve the puzzle concerning heterodimers existence. For example it is known that not al APs from plant origin are heterodimeric (families A1, A3, A11 and A22, Table 1) since many are monomeric and have a broad range of sizes. Some authors believe that the underlying issue in heterodimer formation, at least in what concerns the APs from plant origin is the presence or absence of proteolytic processing enzymes that convert the preproform of the AP into a mature enzyme (Mutlu *et al*., 1999). This could also apply to the rest of APs produced in other organisms.

1.2.6. Stability studies of aspartic proteinases

The genetic code translates almost linearly DNA sequence into an amino acid sequence. Yet, there is no simple "folding code" that translates an amino acid sequence into a three-dimensional protein structure. To understand the basic principles that relate folding, stability and structure the knowledge of the three-dimensional structures of a representative subset of all proteins encoded within a genome is not enough. As important is the study of protein structure thermodynamics and of folding kinetics.

Little has been published about the denatured states of APs, essential for understanding protein stability. Studies have been hampered by the difficulty of purifying large quantities of pure and homogenous protein to perform biophysical and spectroscopic studies (Sarkkinen *et* a*l*., 1992; Veríssimo *et* a*l*., 1996). This is in much due to the typical AP zymogen processing often implicating several steps compromising purification to homogeneity. On the other hand, in general, they are not expressed abundantly affecting the protein recovery yields. To overcome these limitations, many efforts have been carried out to produce mature recombinant forms of several APs that probably justify the late coming of APs structural stabilities studies. The production of functional recombinant forms has also been a challenge due to the required proforms

processing steps to yield mature protein that some are known to involve other, still unknown, proteases.

Most of the available unfolding stability studies were carried out because of their commercial or pharmaceutical importance, such as pepsin and HIV-1 protease or came from the ease of obtaining pure protein samples, such as is the case of other APs studied. Table 1.3 lists the thermodynamic quantities on AP stability studies obtained until recently.

Pepsin along with being the first enzyme whose crystals were examined by X-ray diffraction patterns (for a review, Fruton, 2002), it was also leading the AP stabilities studies. Through calorimetric studies it was showed (Privalov *et* a*l*., 1981) that thermal denaturation of porcine pepsin is a complex process that proceeds by two distinct stages occurring at different temperatures. pH stability studies revealed that pepsin undergoes conformational transition from the native (at acidic pH) to the denatured state in a narrow pH range (between 6 and 7) and almost completely irreversible (Fruton, 1971). This alkaline denaturation results in unfolding of the *N*-terminal lobe and at the same time in structure integrity of the *C*-terminal lobe (Lin *et* a*l*., 1993). Later, it was hypothesised that the observed non-native characteristics of pepsin structure in the alkaline denatured state could have physiological implication. That is to say that some non native states could play a role in the folding of pepsin or its precursor pepsinogen and also in the pepsin transportation to the stomach lumen (Kamatari *et* a*l*., 2003; Campos *et* a*l*., 2003).

In fact, enormous interest in studying the denaturation and stability of non retroviral APs comes from their zymogen derived origin. Several lines of evidence suggest that the structures of zymogen derived proteins in their denatured states have unusual features. Such as that under mild denaturant conditions denatured states can be found, due to the proteolytic processing, and that the cooperativity of the folded state may be reduced (Eder *et* a*l*., 1995; Cunningham *et* a*l*., 1999), as seen with pepsin. Defining the structural characteristics of such proteins should give insights not only into the functional properties of this very important family of enzymes involved in proteolysis, but also in general factors defining protein structure, folding, and activity.

Table 1.3: Gibbs energy change - molar values in aspartic proteinases (adapted and updated from Pfeil, 1998).

Enzyme	Form	pH	T	ΔG (kJ/mol)	Approach	Ref
Pepsin	Free	6.5	25	45.2	DSC (a)	Privalov, 1982
	Inhibited	6.5	25	45.2	DSC (a)	Privalov *et al.*, 1981
	Free	7.5	25	16±1.5	Urea	McPhie, 1989
	Inhibited	7.5	25	17.4±1.8	Urea	Privalov *et al.*, 1981
Rhizopuspepsin	Wild type	3.0	25	66.9±7.1	GuHCl	McPhie, 1989
Wild-type and mutant	P16	3.0	25	71.1±7.5	GuHCl	McPhie, 1989
	Asp30→Ile	3.0	25	45.6±3.3	GuHCl	McPhie, 1989
	Asp77→Thr	3.0	25	63.2±5.9	GuHCl	McPhie, 1989
	Asp30→Ile and Asp77→Thr	3.0	25	46.0±2.9	GuHCl	McPhie, 1989
M. miehei AP	Free	5.4	25	55.3±10.8	GuHCl	Brown *et al.*, 1991
	Free	5.4	25	41.6±0.6	GuHCl	Brown *et al.*, 1991
E. parasitica AP	Free	5.4	25	28.6±1.0	GuHCl	Brown *et al.*, 1991
	Free	5.4	25	22.0±0.5	GuHCl	Brown *et al.*, 1991
	Free	5.4	25	29.8±0.7	GuHCl	Brown *et al.*, 1991
HIV-Protease	Free	6.0	25	59.4±5.9	Urea	Grant et al., 1992
	Inhibited	6.0	25	80.8±2.9	Urea	Brown *et al.*, 1991
	Free	5.0	25	60.67±1	DSC	Todd *et al.*, 1998
	Free	3.4	25	41.84±1	DSC	Todd *et al.*, 1998
HIV-Protease recombinant	0.1 M NaCl	7.0	25	9.02±0.29	Urea	Szeltner *et al.*, 1996
	1.0 M NaCl	7.0	25	13.13±0.53	Urea	Szeltner *et al.*, 1996
	0.1 M NaCl	5.0	25	11.58±0.41	Urea	Szeltner *et al.*, 1996
	1.0 M NaCl	5.0	25	14.22±0.95	Urea	Szeltner *et al.*, 1996
HIV-Protease mutant	0.1 M NaCl	7.0	25	10.40±0.41	Urea	Szeltner *et al.*, 1996
(Gln7→Lys, Leu33→Ile	1.0 M NaCl	7.0	25	13.55±0.52	Urea	Szeltner *et al.*, 1996
Leu63→Ile)	0.1 M NaCl	5.0	25	12.63±0.64	Urea	Szeltner *et al.*, 1996
	1.0 M NaCl	5.0	25	14.35±0.84	Urea	Szeltner *et al.*, 1996
SIV-Protease	Free	6.0	25	55.6±5.4	Urea b)	Grant *et al.*,1992

a) multidomain protein with, 7.9 and 12.5 KJ/mol for the *N*-terminal lobe, 16.3 and 8.4 KJ/mol for the *C*-terminal lobe.
b) ΔG refers o the unfolding of the dimer into monomers, *i.e.*, $N_2 \rightarrow 2U$.

HIV-1, a non zymogen derived enzyme, is also a fine example of an AP that has had its structural stability explored. Design of high affinity transition state analogs that compete with natural substrates for the active site and the development of compounds that inactivate the protease by destabilizing its quaternary or tertiary structure have

been the most common strategies in the development of HIV-1 protease inhibitors (Wlodawer *et al*., 1998). Consequently, an accurate knowledge of the energetics of structural stabilization and binding as well as the identification of the regions in the protease molecule that are critical to stability and function were needed. In this sense the energetics of stabilization of HIV-1 protease has been measured by protein fluorescence and DSC (Grant *et al*., 1992 and Todd *et al*., 1998, respectively) (Table 1.3). The denaturation model found, followed by urea induced protein fluorescence changes, conformed to cases in which protein unfolding and dimer dissociation are concomitant processes where folded monomers do not exist (Grant *et al*., 1992). In the same study, further studies with DSC showed that most of the energy of stabilization is provided by the dimerization interface and that the isolated subunits are intrinsically unstable. The existence of regions of the protease with only marginal stability and a high propensity to undergo independent local unfolding was also predicted (Todd *et al*., 1998). This propensity was related with the ability of some mutantss to obtain resistance towards certain inhibitors.

Thermal denaturation of *Endothia parasitica* and *Mucor miehei* APs was shown to be irreversible, as chymosin (described below in more detail), and unlike to what has been found for the other APs mentioned here (Brown *et al*., 1991). Thermal and guanidine hydrochloride denaturation of these proteinases produced first-order, two-state, kinetic behaviour. Equilibrium unfolding transitions of these proteinases were highly cooperative, also atypical in APs studied so far, like pepsin and HIV-1 protease, nevertheless some deviation from two-state character was found.

The denaturation and renaturation of the single-chain cathepsin D, purified from bovine spleen has been studied by CD, protein fluorescence and enzyme activity (Lah *et al*., 1984). Activity was lost irreversibly on unfolding, but the loss of backbone ellipticity and of folded aromatic environment was partial reversible. The enzyme unfolds in two main stages suggesting the existence of at least two intermediate forms between the native and the fully unfolded states, suggesting no protein folding cooperativity.

Recombinant chymosin was also characterized concerning the thermal stability envisaging large-scale production for the cheese-making industry (Beldarraín *et al*., 2000). It showed irreversible, highly scan-rate-dependent thermal denaturation. At pH 5, the most stable condition, the denaturation could be fitted to the two-state

irreversible model implying the impossibility of obtaining the thermodynamic parameters. Nevertheless, it could be hypothesized that both lobes presented different stabilities, apparently being the *N*-terminal part of the molecule less stable.

Having in mind the number of APs identified so far, there are still few studies concerning the conformational stability of this family. From non monomeric retroviral APs, two apparent different unfolding behaviours can be seen, one where the molecule unfolds cooperatively, and other where the *N*-terminal lobe seems to be less stable, comparing with the *C*-terminal lobe. As for the homodimeric retroviral proteases, instability of the folded monomers was shown to exist.

An integrated discussion comparing stability with structure/function and phylogeny features of the APs characterized so far is still lacking. The APs represent an excellent opportunity to investigate structure and function relationships of proteins owing to the availability of the structural and kinetic data, different degrees of amino acid sequence homology and different specificities among the group members, even though more studies with more members of this family should be encouraged.

1.2.7. Cardosin A

1.2.7.1. General characteristics

Cardosin A was isolated and purified from fresh stigmas of *Cynara cardunculus* L in middle 1990s (Veríssimo *et al*., 1996) and shown to be an heterodimeric molecule composed by two polypeptide chains (apparent molecular weights of 31 and 15 kDa). Inhibition studies showed that it belongs to the APs, with optimum action at pH 4.5 (Sarmento *et al*., 2004b), similar to that of cyprosins and relatively higher in comparison with other plant APs (pH 2-4) (Simões *et al*., 2004). Has seen in some APs, it has two glycosylation sites, one in each polypeptide chain (Asn67 and Asn257) attaching in total 19 sugar rings. They are located at the molecular surface and away from the active site unlikely affecting activity and specificity of the enzyme. It is possible that they protect against accelerated proteolytic cleavage playing therefore a role in the stability

of the molecule as seen with other APs like renin (Hori *et al*., 1988) and with other glycoproteins (Khan *et al*., 2003).

Cardosin A was the first plant APs that has had its three dimensional structure determined (Frazão *et al*., 1999). It is composed by two polypeptide chains in its mature form and presents a very similar fold to family A1 of APs, as can be seen in Figure 1.4. The overall secondary structure consists essentially of *β*-strands with very little *a*-helix. The molecule is bilobal with the active site located in a large cleft between the two identical *β*-barrel like domains each containing one of the catalytic motifs, DTG and DSG both present in the 31 kDa polypeptide chain (Figure 1.4). These catalytic motifs are highly conserved in APs from plants. As in other APs structures, there is a flap that projects out over the cleft and encloses substrates and inhibitors in the active site and where both flap flexibility and certain amino acids, like Tyr75 and Thr77, have been shown to play a role in catalysis (Sielecki *et al*., 1990; Chen *et al*., 1992; Okoniewska *et al*., 1999). Like most APs, this enzyme has a preference to cleave peptide bonds between hydrophobic amino acids and with its primary specificity closely resembling those of cathepsin D and plasmepsins (Castanheira *et al*., 2005). Three intrachains disulfide bridges stabilize the structure, two within the large polypeptide chain (Cys45/Cys56 and Cys206/Cys210) and the third within the small polypeptide chain (Cys249/Cys282). These disulfide bridges positions are highly conserved in the AP family (Davies, 1990). Nevertheless, the two polypeptide chains of cardosin A are held together only by hydrophobic interactions and hydrogen bonds arising with the AP fold (Frazão *et al*., 1999).

Concerning the existence of structural determinants, cardosin A is unique among known plant APs in having an RGD motif located at the surface of the protein belonging to the 31 kDa polypeptide chain (Frazão *et al*., 1999), as can be seen in Figure 1.4. This amino acid sequence in proteins is a well known integrin-binding motif suggesting the involvement of cardosin A in adhesion-dependent recognition events (Faro *et al*., 1990). Furthermore, an additional KGE sequence present in the enzyme has shown to also have a role in this interaction (Simões *et al*., 2005). It is located at the tip of a long loop in the 15 kDa polypeptide chain and is similar to RGD in what concerns charge distribution and location, also at the molecular surface (Figure 1.4).

A

DSGSAVVALT NDRDTSYFGE IGIGTPPQK FTVIFDTGSS VLWVPSSKCI NSKACRAHSM
YESSDSSTYK ENGTFGAIIY GTGSITGFFS QDSVTIGDLV VKEQDFIEAT DEADNVFLHR
LFDGILGLSF QTISVPVWYN MLNQGLVKER RFSFWLNRNV DEEEGGELVF GGLDPNHFRG
DHTYVPVTYQ YYWQFGIGDV LIGDKSTGFC APGCQAFADS GTSLLSGPTA IVTQINHAIG
ANGTSSEELQ VDCNTLSSMN VSFTIGGKKF GLTPEQYILK VGKGEATQCI SGFTAMDATL
LGPLWILGDV FMRPYHTVFD YGNLLVGFAE AA

B

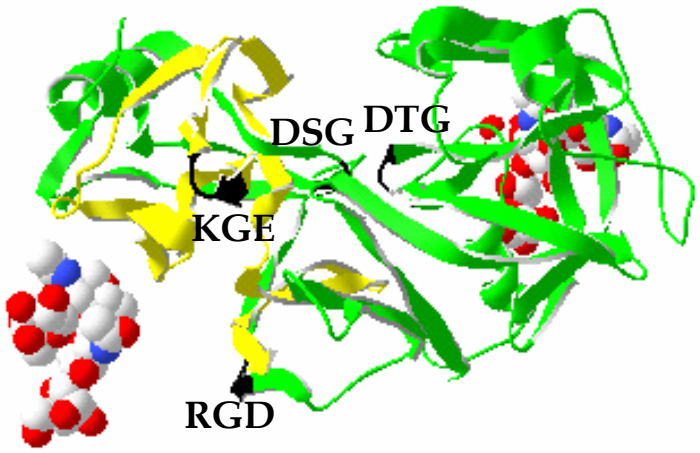

Figure 1.4: Representation of cardosin A. (A) cardosin A amino acid sequence; (B) cardosin A crystal molecule (accession code 1b5f). The polypeptide chains are highlighted in green (31 kDa chain) and yellow (15 kDa chain). The catalytic triads are represented as DSG/DTG (Asp32-Ser-Gly and Asp215-Thr-Gly, respectively). The RGD and KGE motifs (Arg176-Gly-Asp and Lys278-Gly-Glu, respectively) are also represented. Glycosylation is shown in space-filling form.

1.2.7.2. Function, localization and processing

In general, plant APs function is still not known in detail but it is known that they are implicated with death cell events associated with plant senescence, stress responses, programmed cell death and plant sexual reproduction (Simões *et al*., 2004; Radlowski, 2005). Up until now it is believed that cardosin A may be involved primarily in the interaction with pathogens and pollen, that is to say involved in defence and pollen-recognition mechanisms and also involved in the annual senescence of the flower

(Ramalho-Santos *et al*., 1997). Recently an association between cardosin A and phospholipase Dα, known to be involved in degradative cellular processes (Wang, 2000), has been shown to exist, suggesting a possible concerted action in degenerative processes. RGD, and later KGE, motifs have been demonstrated to play essential role in this complex formation (Frazão *et al*., 1999; Simões *et al*., 2005).

Cardosin A is abundantly expressed and limited to cardoon pistils. This vacuolar enzyme is found in papillary layer cells of the stigma and in smaller amounts in the epidermis of the style. Furthermore, transfer of the enzyme from the vacuoles to cell walls and to the apoplast also seems to occur (Faro *et al*., 1990). In vacuoles, it accumulates until the later stages of flower senescence (Ramalho-Santos *et al*., 1997). However, gene expression occurs more importantly in the early stages of flower development.

The primary translation product of cardosin A, the major AP component present in *C. cardunculus* L, consists of a single-chain preproenzyme comprising a signal peptide, a prosegment and the saposin-like domain, PSI, which are all sequentially removed to yield a mature enzyme. It is synthesized and translocated into the rough endoplasmic reticulum as a preproenzyme where it becomes *N*-glycosylated at its glycosylation sites. The processing of newly synthesised cardosin A zymogen starts with the removal of the signal peptide during the passage to the endoplasmic reticulum. The procardosin A is then converted into a two chain enzyme in a multi-step process as the flower matures. Figure 1.5 shows the assumed proteolytic processing of cardosin A.

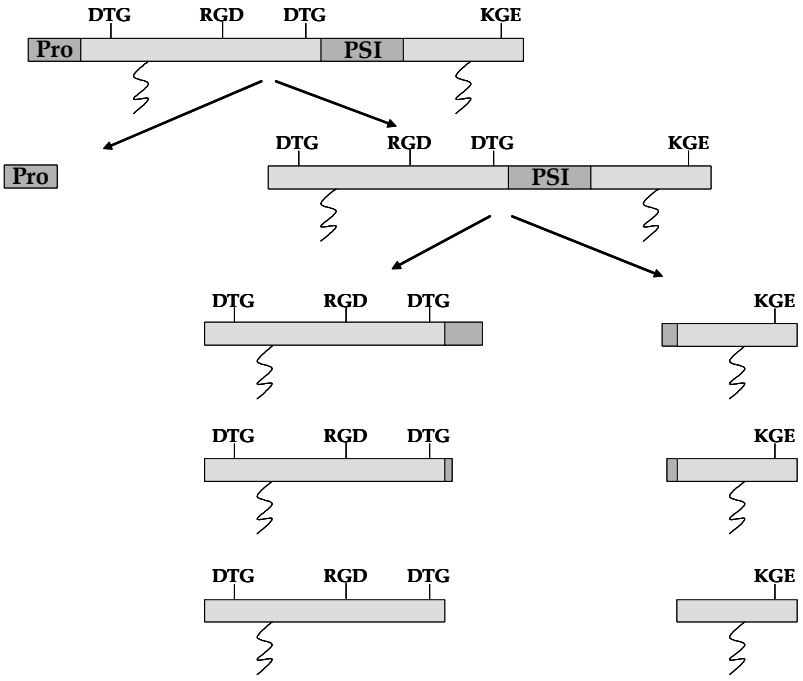

Figure 1.5: Proteolytic processing scheme of procardosin A. Approximate locations of important structure features of cardosin A are signalled: glycosylation sites, catalytic residues in the active site are marked by DTG/DSG, the plant-specific domain is marked as PSI and structural determinants such as RGD and KGE are also highlighted.

The first step in the proteolytic processing is the removal of the propeptide, generating an active intermediate with the PSI present (Ramalho-Santos *et al*., 1998), in a similar way as has been reported for the activation of gastric proteinases (Richter *et al*., 1988). The removal of the propeptide is believed to occur inside the vacuole aided by a decrease in pH. The second step is the removal of the PSI, which is sequentially excised from the precursor (Castanheira *et al*., 2005).

1.2.7.3. Biotechnological applications and structural studies

Cardosin A is present in the pistils of the cardoon *Cynara cardunculus* L. These pistils have been used for centuries in Portugal in the manufacture of *Serra da Estrela* cheese to initiate the milk clotting process. Cardosin A catalytic activity accounted for most events of primary proteolysis (Silva *et al*., 2005). Traditionally, this hydrolytic activity has found practical application in dairy industry.

Since the purification and characterization of cardosin A a lot of attention has been drawn towards this enzyme as seen in the number of publications since the last decade. The fact that it is abundantly expressed in pistils of the cardoon, and that the purification protocol is simple allowing the recovery of several milligrams of pure protein facilitated the characterization of this enzyme and its use for biotechnological applications.

Taking advantage of its specificity requirements contrasting with other used biocatalysts, cardosin A has been studied concerning its potential to synthesise peptide bonds. This enzyme proved to be a good catalyst in the synthesis of peptide bonds with yields comparable with those found for pepsin (Sarmento *et al*, 1998). Besides the possible biotechnological application the numerous reactions carried out allowed a wide study of cardosin A primary and secondary specificity. These results compared with specificity features of other aspartic proteinases improved the knowledge about the particularities of this enzyme active site sub-sites (Sarmento *et al*., 2004; Sarmento, 2002).

Since for hydrolytic reactions peptide synthesis is thermodynamically unfavourable in aqueous buffers, two-phase systems were used (Nakanishi *et al*., 1986) taking advantage of solubility of reaction products in selected organic solvents. The use of organic solvents in reaction synthesis implied that the effects of organic solvents on cardosin A activity, specificity and stability had to be investigated. The investigation showed that cardosin A is stable and active in some organic solvents (Sarmento *et al*., 2003) and later proposed as a reliable probe for limited proteolysis in the presence of organic solvents (Sarmento *et al*., 2006). This could have application in structural characterisation of membrane proteins for surface mapping purposes.

The colagenolytic activity of APs from cardoon pistils, namely cardosin A, has been characterized. It was shown that cardosin A uses a mechanism of collagenolytic

activity similar to the mammalian cathepsin K collagenolytic action (Duarte *et al*., 2005). It was the first report of a plant proteolytic enzyme with collagenolytic activity behaving more closely as animal collagenases on extracellular matrix remodelling. On the other hand, these remodelling processes in animals could be compared with remodelling and/or degradation of the pistil extracellular matrix events during pollen tube growth. In this sense, recently, cardosin A together with the remaining proteases present in the cardoon pistils have been studied concerning their enormous potential for their use in establishing primary cultures of neurons by means of assisting in cell dissociation step during tissue disaggregation (Duarte, 2006; Duarte *et al*., 2006). Also, the colagenolytic activity of cardosin A was tested to act as an antifibrotic agent for post chirurgical abdominal cavity prevention and reduction of adhesions formation (Pereira *et al*., 2005).

New perspectives in biotechnological applications of cardosin A will certainly benefit from further studies of its structure, activity, stability and physiological functions. Meantime the knowledge gained, concerning the cardosin A molecule, will enrich largely AP family knowledge and stimulate structure/function comparisons. The fact that this heterodimeric protein follows the pepsin-like single chain topology makes it an attractive model protein to pursue equilibrium denaturation studies for providing important information on subunit interactions. In reality, understanding the structure-function relationships of an enzyme under different conditions is fundamentally important for both theoretical and applicative aspects. Such studies may provide insight into the molecular basis of the stability of the enzyme, which in turn can be used to design protocols and/or a protein with special properties for biotechnological applications.

3 Objectives

With the proteomics coming of age, it is increasingly important to understand the structure-function relationship of proteins using various biophysical tools (Neet *et al.*, 2002). Cardosin A has revealed to be a fine model protein belonging to the aspartic proteinases family. It has been well characterized in what concerns its structure and enzymatic activity.

The present work intends to perform extensive studies on cardosin A at several pH values and in the presence of acetonitrile. Different spectroscopic methods (e.g., circular dichroism, intrinsic fluorescence and absorbance), activity measurements and calorimetric analysis will be employed to detect and characterize the pH and organic solvent induced states.

Detailed studies of such states can provide useful information for understanding the folding within the APs familiy and the structure-function relationships.

The following questions will be addressed:

Are pH and acetonitrile induced states of cardosin A related to each other?

Do these conformational states furnish any information regarding protein structure and folding in the AP family?

And is there any association between protein stability and folding and the protein oligomeric state?

Does this oligomeric state bring any advantage to the putative physiological roles of cardosin A?

3 Materials and Methods

3.1 Cardosin A purification

Fresh inflorescences were collected from wild cardoons in Ansião, Coimbra and identified as *Cynara cardunculus* L. by Dra. Rosa Pinho (Biology Department Herbarium, University of Aveiro). After harvesting, the pistils were frozen (-20 ^{o}C) until cardosin A purification. Throughout this work pistils harvested in 2000, 2001 and 2002 were used and no differences in cardosin A purification yields and activity were found.

Cardosin A purification was obtained according to a previously optimized process (Sarmento, 2002; Sarmento *et al*., 2004) and described in detail below.

Chromatographies were performed at room temperature in an automated AKTA FPLC (GE Healthcare) equipped with a P-920 pump and a UPC-900 detector, for both ultraviolet wavelength absorption (280 nm) and conductivity monitorization of the eluent. This system was equipped with UNICORN 3.2 software. All solvents were filtered through a 0.2 µm pore membrane and degassed with helium before use.

Figure 3.1 represents the purification of about 5 mg of cardosin A from 2 g of *Cynara cardunculus* L fresh pistils. It can be seen that the purification of cardosin A is carried out with two successive chromatographic protein separations: first, a size exclusion chromatography (SEC) of the acidic pistils extract and after, an anion exchange chromatography (AEC). In SEC, the peak, signalled in Figure 3.1A, contained a mixture of proteins, among them cardosin A, as well as cardosin B (Faro *et al*., 1995) and also isoforms of cardosin A_0 (Lopes, 2003) (Figure 3.1D-lane 2). Subsequently the AEC is carried out to further purify cardosin A from the rest of the proteins. In this way, cardosin A is eluted in one chromatographic peak as signaled in Figure 3.1B corresponding to pure cardosin A, as can be seen by SDS-PAGE analysis (Figure 3.1D-lane 3). Finally, desalting of cardosin A is achieved by replacing salt buffers with water by SEC (Figure 3.1C).

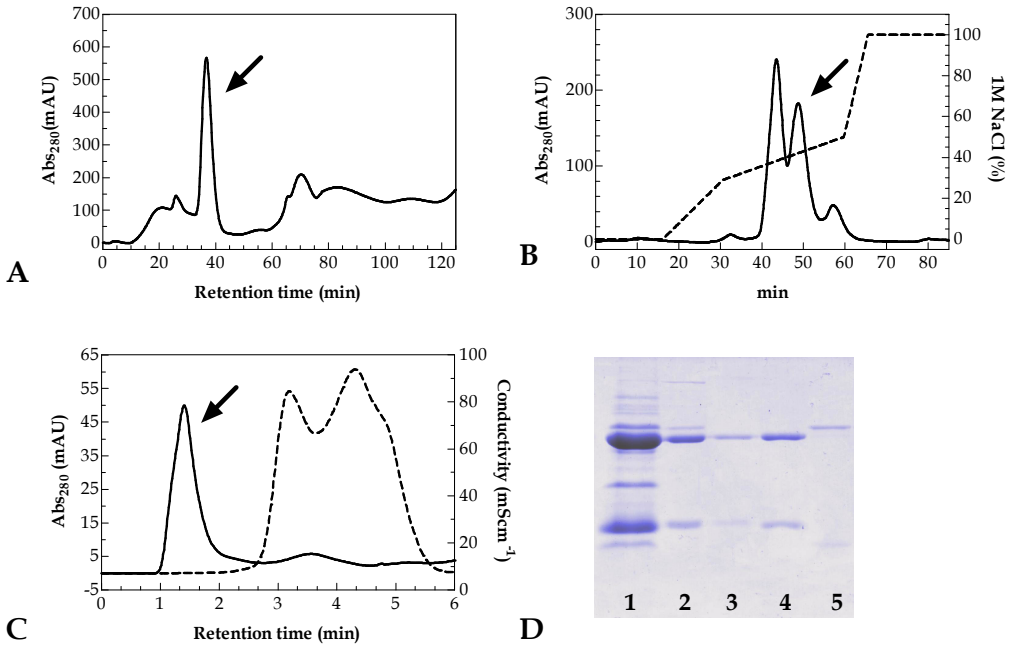

Figure 3.1: Purification of cardosin A from *Cynara cardunculus* L. (A) Size exclusion chromatography (SEC) profile of the pistil's acidic extract monitored at 280 nm. Arrow indicates the chromatographic peak with proteolytic activity; (B) Anion exchange chromatography (AEC) profile of the active faction obtained with SEC. Arrow indicates the chromatographic peak corresponding to pure cardosin A; (C) Desalting chromatographic elution profile of the chromatographic peak fraction corresponding to cardosin A. Desalted cardosin A peak is represented by solid line whereas salts appears as dotted line; (D) Coomassie Brilliant Blue stained SDS-PAGE of the protein factions.. 1- Pistils acidic extract from *Cynara cardunculus* L. 2- Active protein fraction from SEC. 3- Protein faction corresponding to cardosin A obtained by AEC.

Pure cardosin A solution in water was then frozen in liquid Nitrogen and concentrated by lyophilization in a Dura-Dry system for at least 24 hours at a 200 mtorr pressure and a condensator temperature of – 50 ^{o}C.

Lyophilised cardosin A was immediately used or stored at – 20 ^{o}C.

Cardosin A specific activity was determined to be 7.81 ± 0.89 U/mg, with one enzyme unit (U) being the amount of cardosin A needed to hydrolyse 1 μmol of Lys–Pro–Ala–Glu–Phe–Phe(NO_2)–Ala–Leu per minute.

Frozen pistils (2 g) were ground in a mortar with 12 ml of 100 mM sodium citrate buffer (Sigma), pH 3.5 centrifuged at 13000 rpm for 10 minutes. The supernatant was then filtered through a 0.2 μm pore membrane (Schleicher & Schuell).

The filtered supernatant (about 10 ml) was then applied to a Hiload Superdex 75 26/60 semi prep gel filtration column (GE Healthcare). The resin was previously equilibrated with 25 mM Tris.HCl (Amresco), pH 7.6 at a flow rate of 4 ml/min. The active fraction (about 25 ml) was further purified in an anion exchange column, Hiload Q-Sepharose 16/10, previously equilibrated in the same buffer used for size exclusion chromatography. Elution was achieved with a sodium chloride gradient (0.2 – 1 M) at a flow rate of 3 ml/min. Proteolytic fraction containing pure cardosin A was identified by denaturing electrophoresis (see conditions in section 3.3) and stained with Coomassie Brilliant Blue. Salt removal from cardosin A solution was obtained by applying the sample to a HiPrep™ 26/10 desalting (Amersham Pharmacia Biotech) column previously equilibrated in ultra-pure water at a flow rate of 10 ml/min.

3.2 Protein quantification

For activity, fluorescence and chromatographic assays, cardosin A concentration was determined by the BCA™ protein assay (Pierce) according to the manufacturer's instructions. This procedure was used for its ability to quantify dilute protein solutions and for its compatibility in the presence of organic solvents in solution. BSA was used as standard protein. For circular dichroism and calorimetric experiments, cardosin A concentration was determined spectrophotometrically at 280 nm, using a molar absorption coefficient of 43.8×10^3 Mcm $^{-1}$ and considering a 42 kDa molecular weight for the cardosin A heterodimer. For the synthetic peptide, a molar absorption coefficient of 1480 mM cm^{-1} at 300 nm was used (Veríssimo *et al.*, 1996).

3.3 Electrophoretic analysis of cardosin A

Whenever necessary, cardosin A electrophoretic analysis was made under both denaturing and native conditions. Electrophoreses were carried out in a Mini-Protean 3 Cell (Bio-Rad) device linked to a power supply equipment, PowerPac 300 (Bio-Rad).

3.3.1 Denaturing electrophoresis

Sodium dodecyl sulphate polyacrilamide gel electrophoresis (SDS-PAGE) was used to analyse cardosin A solutions purity as well as for analysis of SEC protein samples content. A discontinuous methodology was adopted in what concerns acrylamide percentage, ionic composition and pH of the stacking and running gel (Laemmli, 1970). Optimum acrylamide concentration in stacking gel (15 %) (v/v) (Makowski *et al.*, 1997) was chosen according to the cardosin A molecular weight.

Running gel was prepared by mixing 2.90 ml of distilled water, 3.335 ml of 1.5 M Tris at pH 8.8, 3.75 ml of acrylamide:bisacrylamide (40 %, w/v), 200 µl of SDS (10 %, w/v) and 100 µl of ammonium persulphate (10%, w/v). Polymerisation was initiated by adding 5 µl TEMED. Stacking gel was prepared by mixing 3.260 ml of distilled water, 1.25 ml of 0.625 M Tris at pH 6.8, 488 µl of acrylamide:bisacrylamide (40 %, w/v), 100 µl of SDS (10 %, w/v) and 50 µl of ammonium persulphate (10%, w/v). Polimerization was initiated by addition of 2.5 µl of TEMED. Samples were diluted (50 %, v/v) with denaturant solution and heated at 100 ^{o}C for 3 minutes. Denaturing solution was prepared by mixing β-mercaptoethanol (2 %, v/v), SDS (2 %, w/v), urea (8M) and Tris:Bicine (100mM, 1:1).

Running buffer was prepared by addition of 100 mM Tris, 100 mM Bicine and 0.1 % SDS (w/v).

Gels ran for 60 minutes at 150 V.

3.3.2 Native electrophoresis

Native-PAGE was used to analyse cardosin A molecular associations populations.

An uniform polyacrilamide concentration was used and PAGE was carried out in the absence of SDS.

Gel was prepared by mixing 2.90 ml of distilled water, 3.335 ml of 1.5 M Tris at pH 8.8, 3.75 ml of acrylamide:bisacrylamide (40 %, w/v) and 100 µl of ammonium persulphate (10 %, w/v). Polymerisation was initiated by adding 5 µl TEMED. Protein samples were diluted (50 %, v/v) in glycerol (50 %, w/v). Running buffer was prepared by mixing 100 mM Tris and 100 mM Bicine.

Gels ran 260 minutes at 20 mA.

3.3.3 Gel staining

Whenever necessary, Coomassie Brilliant Blue R-250 (Sigma) or silver nitrate was used.

3.3.3.1 *Coomassie Brilliant* Blue staining

Gels were incubated with slight shaking in staining solution for 1 h at room temperature or until a blue and opaque colouring of the gels is obtained. Destaining of the gels is required to remove the background stain. The gel is soaked in destaining solution, also with slight shaking, until the background is cleared.

Staining solution: 0.4 % (w/v) *Coomassie Brilliant Blue*, 50 % (v/v) methanol and 10 % (v/v) acetic acid.

Destaining solution: 25 % (v/v) and 5% acetic acid (v/v).

3.3.3.2 Silver staining

For detection of low concentration protein bands (in the range of few ng) in polyacrilamide gels, silver staining was used and performed at room temperature according to Wray *et al*. (1981). All steps were done with slight shaking.

Gels were initially fixated in 50 % methanol (v/v), 10 % acetic acid (v/v) for 30 min and then in 5 % methanol (v/v), 7.5 % acetic acid (v/v) for another 30 min.

After, the gels were soaked for 60 min in 7 % glutaraldehyde (v/v) and rinsed in water for 5 minutes.

The gels are then placed in staining solution containing 0.2 % silver nitrate (w/v), 0.25 % ammonia (v/v) for 25 min (in the dark). Afterwards, the gels were rinsed in water for 5 min to remove excess stainer.

For the development step the gels were soaked in 1.85 % formaldehyde (v/v) and 0.25 % citric acid (w/v). This solution reduces silver ions to metallic silver linked to the protein resulting in development of the protein bands. The deposition was allowed to continue until protein bands became visible and the background clear. To end this reaction, the gel was washed in 1 % acetic acid (v/v) that complexes with any free silver to prevent further reduction.

3.4 Enzymatic assays

The proteolytic activity of cardosin A was determined using the synthetic peptide Lys-Pro-Ala-Glu-Phe-Phe (NO_2)-Ala-Leu as substrate and as described before (Sarmento, 2001; Sarmento *et al*., 2007), where the rate of hydrolysis was followed by RP-HPLC.

An AKTA FPLC (GE Healthcare) equipped with a P-920 pump and a UPC-900 detector was used. Additionally, automatic injection of the aliquots was employed (Autosampler A-900). This system was equipped with UNICORN 3.2 software. All solvents were degassed with helium before use.

Assays were performed in a 300 µl reaction mixture containing 0.217 mM of the synthetic peptide in 50 mM sodium acetate buffer with 200 mM sodium chloride and 4 % DMSO (dimethyl sulphoxide) (v/v), pH 4.7. Experiments were carried out at 25 ^{o}C. Enzymatic reaction was started by addition of 0.146 µg cardosin A (final cardosin A concentration was 0.115 x 10-4 mM) and at selected times, aliquots (60 µl) were taken and reaction stopped by addition of 540 µl of 1.5 % TFA (v/v). Samples were automatically injected to a C18 column equilibrated with 0.1 % TFA (v/v) and eluted with an acetonitrile gradient acidified with 0.1 % TFA (v/v). Hydrolysis products were detected at 254 nm.

Substrate hydrolysis velocities were calculated (mol min^{-1}). A calibration curve was built, where, after complete substrate hydrolysis peak areas were correlated with substrate concentration (Sarmento *et al*., 2004).

3.4.1 pH effect on cardosin A activity

Cardosin A activity was determined as a function of pH as previously reported (Oliveira, 2001).

Assays were carried out in a 300 µl reaction mixture containing 0.217 mM of the synthetic peptide in 50 mM CH_3CO_2Na buffer with 200 mM NaCl and 4 % $(Me)_2$ SO, pH 1-9 and carried out as described in 3.4. Solutions were prepared immediately before activity assays were initiated. No pH variation was detected throughout the experiments.

3.4.2 Measurement of cardosin A activity in acetontrile

Cardosin A activity in acetonitrile was studied from two viewpoints: after 1 h incubation in the presence of increasing concentrations of acetonitrile and also time dependent activity changes in 10 % acetonitrile solutions (v/v).

All the reactions were conducted in triplicates and at 25 ºC. No protein or salts precipitation were observed in aqueous-solvent mixtures up to 90 % of acetonitrile (v/v). For time dependent activity changes studies, measurements in buffer (0 % acetonitrile) without incubation were considered as control.

Reversibility of the acetonitrile induced effects on cardosin A activity was also investigated. At least 100 fold dilutions in buffer were assured.

3.4.2.1 Acetonitrile induced effects in cardosin A activity

Cardosin A was allowed to incubate in solutions with acetonitrile for 1h. Incubation buffer was 50 mM sodium acetate buffer with 200 mM sodium chloride, pH 4.7 with varying acetonitrile concentrations (0 – 90 % v/v).

Cardosin A solutions (34.7 x 10-4, 6.95 x 10-4 and 1.39 x 10-4 mM) were incubated for 1 h at 25ºC in incubation buffer and in the presence of varying concentrations of acetonitrile. Reactions were started by addition of cardosin A solution to the reaction media (final concentration 0.115 x 10-4 mM), according to 3.4 and with the same acetonitrile concentration.

3.4.2.2 Cardosin A reactivation experiments

Reactivation assays were conducted by previously incubating (1h) cardosin A in the same acetonitrile concentrations tested above and diluted (about 200 fold) by addition of incubation buffer. Cardosin A was allowed to incubate for 1 h at 25 ºC and assayed for activity as described in 3.4.

Cardosin A (27.8 x 10-4 mM) was incubated for 1 h in incubation buffer and in the presence of varying concentrations of acetonitrile. After this cardosin A solutions were diluted 200 fold by addition of incubation buffer, without acetonitrile, and incubated for another 1 h at 25 ºC.

Reactions started by addition of cardosin A to the reaction media (final cardosin A concentration was 0.115 x 10-4 mM) and performed according to 3.4 and without acetonitrile in the reaction buffer.

3.4.2.3 Cardosin A time dependent activity changes in 10 % acetonitrile

---

Cardosin A was solubilised in incubation buffer (final concentration 6.95 x 10-4 mM) with varying concentrations of acetonitrile and incubated at 25 ºC. At selected times, cardosin A aliquots were taken and enzymatic reactions started immediately by addition of the enzyme to the reaction media, as described in Section 3.4 and with the same amount of acetonitrile.

3.4.2.4 kinetic parameters of Cardosin A in the presence of acetonitrile

Cardosin A follows first-order kinetics, and the kinetic parameters were calculated according to the *Michaelis-Menten* equation (Fersht, 1999). K_m, V_{max} and K_{cat} were determined by measuring initial rates of Lys-Pro-Ala-Glu-Phe-Phe(NO_2)-Ala-Leu hydrolysis at several substrate concentrations.

Assays were performed in a 300 µl reaction mixture containing 0.044 - 1.034 mM of the synthetic peptide Lys-Pro-Ala-Glu-Phe-Phe(NO2)-Ala-Leu in 50 mM sodium acetate buffer with 200 mM sodium chloride and 4 % DMSO, pH 4.7 with different acetonitrile concentrations.

Cardosin A solutions (6.95 x 10-4 mM) were incubated for 1h in incubation buffer and with the appropriate acetonitrile concentrations. No protein or salts precipitation were observed in aqueous-solvent mixtures up to 90% of acetonitrile. Reactions were started by addition of cardosin A aliquots.

Kinetic parameters, K_{cat} and K_m were estimated according to the *Lineweaver-Burk* equation (Lineweaver *et al*., 1934).

3.5 Gel filtration studies

Acetonitrile, pH and temperature induced unfolded states of cardosin A were analyzed by size exclusion chromatography (SEC) using a Superdex 75 HR 10/30 FPLC column (GE Healthcare) and a 100 µl sample injection loop. All measurements were made at room temperature.

For pH induced unfolding studies the mobile phase was 10 mM sodium phosphate buffer, pH 4-13. Samples were eluted at a flow rate of 1 ml/min. For acetonitrile induced unfolding studies, the mobile phase was 10 mM sodium phosphate buffer, pH 5 with the required acetonitrile concentrations.

The protein elution profile was monitored by recording of the absorbance at 280 nm on an AKTA Basic system (GE Healthcare). All solvents were filtered (0.2 µm) and degassed with Helium prior to use.

Each experiment was repeated at least twice to evaluate reproducibility of the chromatographic profiles.

3.5.1 Cardosin A pH induced unfolding

Cardosin A in 10 mM sodium phosphate solutions (2.38 x 10-2 mM) at different pH were prepared immediately before sample injection into the column, previously equilibrated in the same buffer were cardosin A was initially solubilized. No pH variation was detected throughout the experiments.

3.5.2 Cardosin A temperature induced unfolding

Cardosin A in 10 mM sodium phosphate (2.38 x 10-2 mM), pH 5 was heated up to 80 ^{o}C, just below the transition temperature of cardosin A and heated also at 100 ^{o}C and allowed to cool in ice. Cardosin A samples were then placed at room temperature and applied to the column.

3.5.3 Acetonitrile induced unfolding

During sample preparation no protein or salts precipitation were observed in aqueous-solvent mixtures up to 40 % of acetonitrile.

Cardosin A was solubilized in 10 mM sodium phosphate (2.38 x 10^{-2} mM), pH 5 with several acetonitrile concentrations and incubated for 1 h at 25 ºC. After incubation samples were applied to the column previously equilibrated in the same buffer containing the appropriate acetonitrile concentration.

3.5.4 10 % acetonitrile time dependent changes

Cardosin A was solubilized in 10 mM sodium phosphate (1 mg/ml), 10 % acetonitrile, pH 5 and incubated up to 70 h, at 25 ºC. At selected incubation times samples were taken and applied to the column previously equilibrated in the same incubation buffer.

3.5.5 Data analysis, apparent molecular weight estimates

The apparent molecular weights of cardosin A were estimated from a semilogarithmic plot of the M_r values for the calibration proteins and elution volume. A Gel Filtration LMW Calibration Kit (GE Healthcare) was used according to the manufacturers instructions. Blue dextran was used to determine the column void volume (runs at 2000 kDa). The protein standards (13.7, 25, 43 and 67 kDa) were applied individually to the column.

Calibration proteins were solubilized (1 mg/ml) in 100 mM sodium phosphate, pH 5.0 and were applied to the column previously equilibrated in the same buffer.

The relative elution volume was compared to that of molecular mass (M) standards. The relative elution volume was calculated as:

$$K_{av} = \frac{V_{e} \cdot V_0}{V_{g} \cdot V0} \qquad (1)$$

where V_e is the elution volume, V_o is the void volume determined by the elution of blue dextran 2000, and V_g is the geometric column volume. A standard curve was plotted of K_{av} versus *log* (M) (Whitaker, 1963).

3.6 Fluorescence studies of cardosin A

The intrinsic fluorescence of the proteins arises from the aromatic amino acid residues and can be used to monitor global tertiary structure changes (Burstein *et al*., 1983; Schmid, 1997). In this work, the fluorescence of tryptophan residues available in cardosin A was used.

Cardosin A primary structure highlights four tryptophan residues in the 31 kDa polypeptide chain and one in the 15 kDa chain, at positions 39, 137, 154, 190 and 298, respectively, according to pepsin numbering (Frazão *et al*., 1999).

Tryptophan emission fluorescence experiments were done according to previous studies with cardosin A in aqueous buffer (Oliveira, 2001).

Steady-state fluorescence measurements were carried out on a F-4010 Hitachi and on a Jasco Spectrofluorometer FP 770 spectrofluorometers. A fluorescence excitation wavelength of 295 nm was used to avoid the contribution of the emission of residues other than tryptophan. The monochromator slit width was kept at 5 nm in the excitation and emission channels. Fluorescence was measured in the range 300 - 400 nm.

Protein incubations were carried out at 25 ºC and fluorescence measurements at room temperature. 0.5 cm and 1 cm path-length quartz cells were used and solutions constantly stirred.

Fluorescence measurements were carried out on protein solutions with an optical density of less than 0.2 at 295 nm to avoid the inner filter effect.

3.6.1 pH dependence of the emission fluorescence spectra

Cardosin A emission fluorescence spectra study as a function of pH was carried out as reported previously (Oliveira, 2001).

No pH variation was detected throughout the experiments.

Cardosin A was solubilized (12 x 10-4 mM) in 10 mM sodium phosphate at different pH values and spectra collected.

3.6.2 Temperature dependence on the emission fluorescence spectra

The temperature-dependence of cardosin A emission fluorescence spectra was investigated using thermostatically controlled water circulating in a hollow brass cell-holder.

Also, the extent of aggregation of cardosin A solution at pH 5 was measured as a function of temperature by monitoring light scattering at 365 nm (excitation and emission band passes set at 1.5 nm).

The heating rate was between 0.6 and 1.2 ^{o}C min^{-1}, and spectra were collected at the desired temperatures over the entire temperature range.

The temperature of the sample was monitored with a thermocouple immersed in the cell under observation.

Cardosin A was solubilized (4.76 x 10-4 mM) in 10 mM phosphate, pH 5 and spectra immediately collected.

3.6.3 Effect of acetonitrile on cardosin A emission fluorescence

Acetonitrile induced effects in cardosin A were measured by fluorescence after 1 h time incubation.

Since that both the intensity and wavelength of the maximum emission (λ_{max}) of tryptophan fluorescence depend on the solvent polarity, when studying fluorescence at different solvent compositions both these effects were considered. This was done by

comparing enzyme fluorescence with the fluorescence of model tryptophan derivatives such as *N*-acetil-L-tryptophan ethyl ester (ATEE) (Sato *et al*., 2000).

Cardosin A (4.76 x 10-4 mM) was solubilized in 10 mM sodium phosphate, pH 5 in the presence of varying acetonitrile concentrations and incubated for 1 h at 25 ºC. After this, spectra were recorded.

Also small ATEE quantities were solubilized in varying acetonitrile concentrations in 10 mM sodium phosphate buffer solutions, pH 5 and spectra collected.

3.6.4 Cardosin A reversibility of acetonitrile induced emission effects

Refolding studies were conducted by previously incubating (1h) cardosin A in some acetonitrile concentrations and diluted (about 200 fold) by addition of incubation buffer.

Cardosin A was solubilized (47.6 x 10-4 mM) in 10 mM sodium phosphate, pH 5 with varying acetonitrile concentrations and incubated for 1 h. Then, 10 mM sodium phosphate buffer was added to reach 200 fold dilutions, and after 1 h incubation at 25 ºC, spectra were collected.

3.6.5 Time dependent changes of cardosin A emission fluorescence spectra at 10 % acetonitrile

The time-dependent cardosin A conformational changes induced by the presence of 10 % acetonitrile were investigated by recording several emission spectra at different time intervals, ranging from 2 min to 170 h.

Cardosin A (4.76 x 10-4 mM) was solubilized in 10 mM sodium phosphate buffer with 10 % acetonitrile (v/v), pH 5.0. Cardosin A solutions and incubated up to 170 h. At selected times samples were taken and spectra collected.

All spectra were collected at room temperature under continuous stirring and incubations were at 25 ^{o}C.

3.6.6 Data analysis

All fluorescence spectra were corrected for the background fluorescence of the corresponding buffer solution. Both fluorescence intensity, at single excitation and emission wavelengths, as well as the maximum emission wavelengths (λ_{max}) were considered. The position of the middle of the chord drawn at the 80% level of maximum intensity (λ_{max}) was taken as the position of the spectrum (Oliveira, 2001).

At least three spectra were recorded for each experiment and then averaged.

Whenever necessary, the fluorescence spectra of cardosin A were analysed on the basis of the model of discrete sates of tryptophan residues in proteins. This model reflects the existence in proteins of five statistically most probable classes of discrete states of tryptophan residues (classes A, S, I, II and III) ranging from deep buried tryptophan residues to fully exposed and surrounded by solvent (Burstein, 1973). This is based in the assumption that tryptophan residues are located in a few kinds of preferred environments in protein structures that provide different combinations of specific interactions of excited fluorophore with the environment. To effectively decompose tryptophan fluorescence spectra of proteins into individual and identified components, algorithms have been developed (Burstein *et al*., 1973; Reshetnyak *et al*., 2001). In this study the modified SIMS algorithm was applied (Burstein *et al*., 2001).

Briefly, all the fluorescence spectra were normalized before calculation. All classes of Trp residues states were assumed to be present in the protein and were included in the curve-fitting process. A fitting function was used to describe the possible components of a spectrum. All possible fittings were obtained by varying the maximum intensity (I_m) and wavelength (V_m) of each component independently within appropriate limits

and the best fitting results were obtained according to the least root mean-square criterion.

3.7 Circular dichroism studies

Circular dichroism (CD) has been shown to be very sensitive to the secondary structure of polypeptides and proteins. CD spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light (rather than the commonly used absorbance) by a substance. It has been shown that CD spectra between 260 and approximately 180 nm can be analyzed for the different secondary structural type contents: alpha helix, parallel and antiparallel beta sheet, turn, and other (Venyaminov *et al*., 1996).

Circular dichroism experiments were recorded on a Jasco-715 spectropolarimeter, using a spectral band-pass of 2 nm and a cell path-length of 1 mm under constant nitrogen flow. In a far-UV region (195-260 nm) spectra were recorded in a 0.1 cm cell and with a scan speed of 50 $nm.min^{-1}$, a 2 nm bandwidth and 1 s integration time.

Protein incubations were always carried out at 25 ^{o}C and circular dichroism spectra collected at 25 ^{o}C and also at room temperature.

3.7.1 Ellipticity changes of cardosin A as a function of pH

Cardosin A circular dichroism study as a function of pH was carried according to Oliveira (2001). No pH variation was detected throughout the experiments.

Cardosin A (ranging from 2.38 x 10-4 to 47.6 x 10-4 mM) in 10 mM sodium phosphate at different pH were prepared immediately before spectra were collected.

3.7.2 Temperature dependence on ellipticity changes

The dependence of cardosin A ellipticity on temperature at pH 5 was investigated by heating the protein solution from 10 to 90 ºC at a constant heating rate (1 ºC/min), using a NESLab RT-11 programmable water bath.

Cardosin A was solubilized (from 2.38 x 10-4 to 47.6 x 10-4 mM) in 10 mM sodium phosphate, pH 5 at 25 ºC and spectra immediately collected.

3.7.3 Acetonitrile dependent ellipticity changes

Acetonitrile induced effects in cardosin A were measured by circular dichroism after 1 h of incubation.
No protein or salts precipitation were observed in aqueous-solvent mixtures up to 40 % of acetonitrile.

Cardosin A (concentrations raging from 2.38 x 10-4 to 47.6 x 10-4 mM) was solubilized in 10 mM sodium phosphate, pH 5.0 in the presence of varying acetonitrile concentrations and incubated for 1 h at 25 ºC. After this, spectra were recorded.

Additionally, buffer with varying acetonitrile concentration solutions without cardosin A were prepared and spectra collected.

3.7.4 Time dependent effects of acetonitrile on cardosin A ellipticity

The time-dependent influence of 10 % acetonitrile on the secondary structure of cardosin A was investigated. Several spectra were recorded for each sample at different time intervals, ranging from 2 min to 120 h.

Cardosin A (concentrations raging from 2.38 x 10-4 to 47.6 x 10-4 mM) was solubilized in 10 mM sodium phosphate buffer with 10 % acetonitrile, pH 5. Cardosin A solutions were then incubated up to 170 h. At selected times samples were taken and spectra collected.

Spectra were collected at room temperature and cardosin A incubations were carried out at 25 ^{o}C.

3.7.5 Data analysis

Spectra were measured 4 times, averaged and corrected by subtraction of the solvent spectrum obtained under similar conditions.

Results are expressed in terms of molar ellipticity [Θ] and calculated by the following equation:

$$[\Theta] = \frac{\Theta_{obs} \times 100 \times M_{res}}{c \times l} \qquad (2)$$

where, Θ_{obs} is the ellipticity (degrees) measured at wavelength (λ), M_{res} is the mean residue molar mass (protein molecular mass divided by the number of amino acid residues in the protein), c is the protein concentration (mg/ml) and l is the optical path-length of the cell (dm).

3.8 Differential scanning calorimetry studies

Solution differential scanning calorimetry (DSC) measures the difference in heat energy uptake between a sample solution and appropriate reference (buffer/solvent) with increase in temperature. It has been widely employed during the last two decades to study thermal transitions in proteins (Bruylants, 2005). It can provide detailed information on the energetics and mechanism of the folding/unfolding processes of proteins (Cooper, 1998).

The calorimetric experiments were performed on a MicroCal MC-2D differential scanning microcalorimeter (MicroCal Inc., Northampton, MA) with cell volume of 1.22 ml. The calorimetric unit was interfaced with an IBM-compatible computer for automatic data collection and analysis.

The samples and reference solutions were properly degassed in an evacuated chamber at room temperature and carefully loaded into the calorimeter to eliminate bubbling effects. An overpressure of 2 atm of dry hydrogen was kept over the liquids in the cells to prevent any degassing during heating.

Overall, unless noted, heating scan rate of 60 ^{o}C/h was used in these experiments, since no significant scan rate effect was observable in all results shown.

The baseline was obtained with buffer in both reference and sample cells.

To test for renaturation, the sample was cooled at the end of the first scan, allowed to re-equilibrate to the starting temperature, and then scanned again. The percentage of renaturation is expressed as the calorimetric enthalpy change (ΔH_{cal}) of the second scan divided by that of the first one.

3.8.1 Thermostability of cardosin A in aqueous systems

Cardosin A (concentrations from 170 x 10-4 to 300 x 10-4 mM) was solubilized in 10 mM citrate phosphate, at different pH values, at 25 ^{o}C. After, the samples were degassed as explained above and loaded into the calorimeter sample cell and heating scan initiated and data collected.

3.8.2 Thermostability of cardosin A in 10 % acetonitrile

The time-dependent influence of 10 % acetonitrile on the thermostability of cardosin A was investigated by recording several scans for each sample at different time intervals, ranging from 0 h to 71 h.

Cardosin A (concentrations from 170 x 10-4 to 300 x 10-4 mM) was solubilized in 10 mM citrate phosphate buffer with 10 % acetonitrile, pH 5. Cardosin A solutions were then incubated up to 71 h at 25 ºC. At selected times, samples were taken and degassed as explained above and loaded into the calorimeter sample cell. Heating scans were immediately initiated and data collected.

For acetonitrile experiments, the baseline was obtained with 10 % acetonitrile in both reference and sample cells.

3.8.3 Data analysis

The molar excess heat capacity curves were obtained by baseline subtraction and normalization with the protein concentrations and volume of the calorimeter cell. After, the curves were smoothed and plotted using the Windows-based software package (ORIGIN) supplied by MicroCal.

Data were analyzed by non-linear regression statistical fitting using the two-state folding/unfolding model both for the investigation of cardosin A thermostability in aqueous system and in aqueous system with 10 % acetonitrile (Takahashi *et al*., 1981).

4 Results and Discussion

4.1 The effect of pH

The structure and function of most macromolecules are affected by pH, and most proteins operate optimally at a particular pH. This effect can be noticed at two levels. First the pH can have an effect on the state of ionization of acidic or basic amino acids, altering the ionic bonds that help to determine the three dimensional shape of the protein, causing therefore conformational changes. This can lead to altered protein recognition or enzyme inactivation. Furthermore, alterations of the ionization state of catalytically important amino acids, such as Asp215 and Asp32 in pepsin-like APs, can preclude catalysis. The stabilization of native enzyme structures is intuitively important to permit the catalytic function. Monitoring denaturing induced effects on both structure and function of enzymes allows picturing the underlying structure function interactions.

4.1.1 Dissection of pH induced effects on cardosin A

Cardosin A pH optimum for catalytic activity, at 25 ^{o}C, was shown to be 4.5 using the synthetic substrate Lys-Pro-Ala-Glu-Phe-Phe(NO_2)-Ala-Leu (Sarmento *et al*., 2004b). The pH dependence curve is bell shaped and the values for the active site ionization constants, pK_{a1} and pK_{a2} of the free enzyme are known, 2.5 ± 0.2 and 5.3 ± 0.2 (Veríssimo *et al*., 1996), respectively, falling in the range of the active site ionization constants determined for other APs, as seen in Table 4.1.

Table 4.1: pH related features of some aspartic proteinases.

Enzyme	pH optimum	pK_{a1}	pK_{a2}	Location	Ratio[a]	Ref.
Cardosin A	4.5 (2.5-7.5)	2.5±0.2	5.3±0.2	Protein storage vacuoles	1.48	Veríssimo *et al.*, 1996
Pepsin A	1-2 (1-5)	1.57±0.04	5.02±0.05	Acidic lumen of stomach	8.75	Lin *et al.*, 1992
HIV1-PR	5.5-7.5	3.32±0.1	6.80±0.1	Cytosol	0.73	Ido *et al.*, 1991

a) Ratio between acidic and basic amino acids in the protein.

4.1.1.1 pH induced unfolding – Structure function analysis

In a previous investigation cardosin A was already studied concerning its pH induced effects on activity, secondary and tertiary structures and results briefly discussed (Oliveira, 2001). Figure 4.1A and B, shows activity, fluorescence and circular dichroism results obtained therein. Overall it was seen that cardosin A, at 25 ^{o}C, is active in the pH range 2.5-7.5, with maximum activity at 4.5, presenting the typical bell shaped curve associated from the summation of the curves for the two catalytic aspartate titrations (James *et al.*, 1985). In this pH range, cardosin A displays identical far UV CD and fluorescence spectra supporting that the pH induced inactivation is not related with the pH induced unfolding of the protein structure. Below and above this range, at highly acidic and alkaline conditions, changes in secondary and tertiary structure were detected.

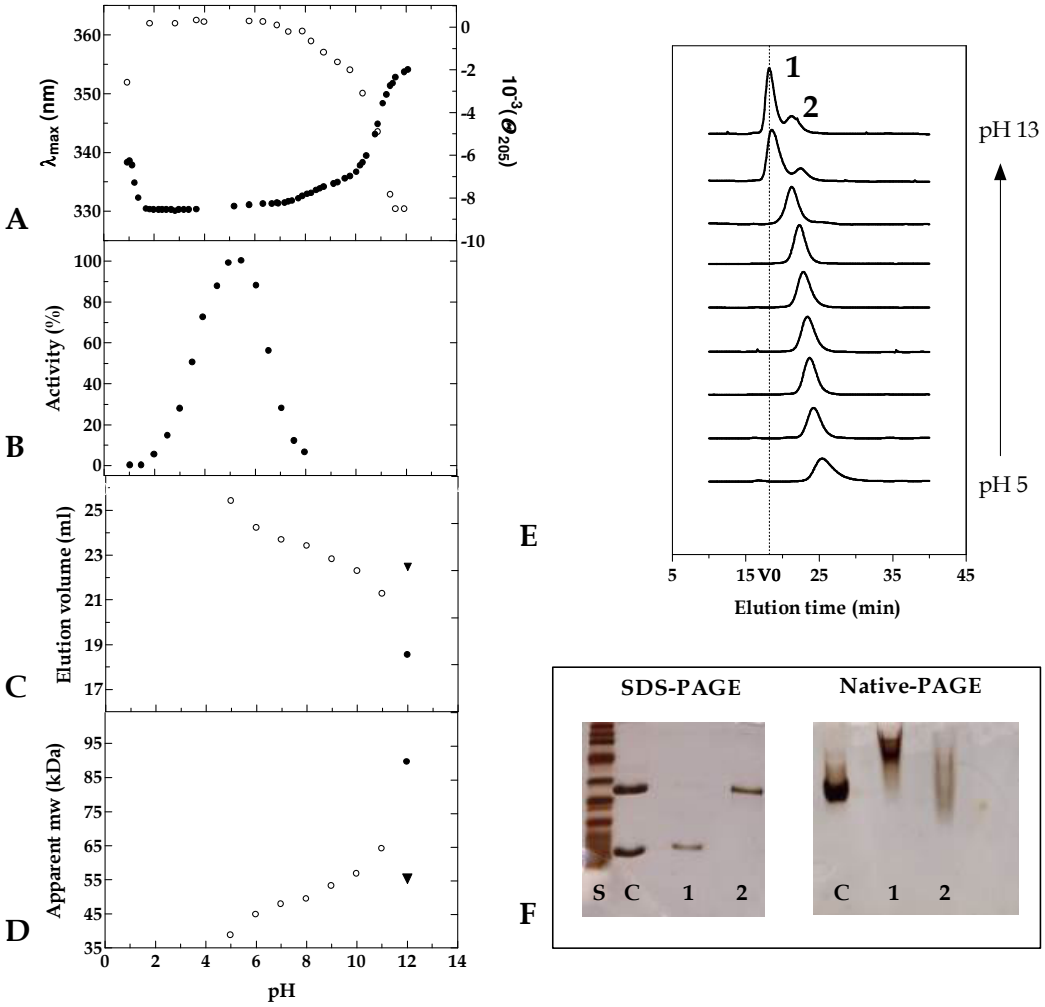

Figure 4.1: Structural and activity changes of cardosin A with varying pH. (A) Fluorescence spectrum position (closed circles) and ellipticity at 205 nm (open circles) (adapted from Oliveira, 2001); (B) Enzymatic activity (closed circles) (adapted from Oliveira, 2001); (C) Elution volume by size exclusion chromatography of cardosin A at different pH; (D) Cardosin A apparent molecular weight at different pH by extrapolation of elution volumes using protein molecular weight calibration kit. For (C) and (D), open and closed circles and triangles stand for major elution peaks (sections 3.5 and 3.5.1); (E) Size-exclusion chromatographic elution profiles and electrophoretic analysis of cardosin A in aqueous buffer at different pH; (F) Electrophoretic analysis of peaks 1 and 2 obtained from (E). Standard molecular weights (S) are, from top of the lane to bottom: 66, 45, 34.7, 24, 18.4 and 14.3 kDa. Cardosin A was used as control, named C in the figure. Size exclusion chromatographies were carried out at room temperature according to sections 3.5 and 3.5.5. Electrophoretic analyses were carried out according to section 3.3.

Overall, below pH 2.5, where cardosin A is inactivated, a sharp red shift of the fluorescence spectrum as well as an increase of the negative value of ellipticity at 205 nm are seen indicating simultaneous exposure of tryptophan residues and secondary structure destabilization. Also, at pH below 2.5 activity is ruled out probably since the formation and stabilization by hydrogen bonds of the tetrahedral intermediate is not favored with the protonated Asp 32 (pK_a = 2.5 ± 0.2). At the same time other carboxyl groups can also be protonated. All together this could result in unfavorable electrostatic interactions introduced by the increase in positive charge on cardosin A molecule inducing unfolding as observed by shifts in fluorescence and CD data.

Clearly the acid and alkaline induced denatured states of cardosin A are different as seen by the spectroscopic data, where both the maximum fluorescence emission and the ellipticity at 205 nm are dissimilar, suggesting a monophasic acid denaturation and a biphasic alkaline denaturation transition as described below.

At pH above 7.5, cardosin A is inactivated, probably due to deprotonation of Asp215 (pK_a = 5.3±0.2). Also, a two-step red shift in the fluorescence spectrum and a coordinated two-step change in ellipticity at 205 nm were observed. The alkaline denaturation of cardosin A is possibly triggered by deprotonation of some Cys (pK_a around 8), Tyr and Lys (pK_a around 10) residues. An analysis of the amino acid sequence of cardosin A (Figure 4.2) shows that there are 5 Cys, 7 Lys and 9 Tyr residues in the large chain and 2 Cys, 4 Lys and 3 Tyr residues in the small chain, demonstrating that deprotonation of residues at alkaline conditions will likely occur. The second subsequent and sharper transition takes place, registered both by fluorescence and CD resulting in an alkaline denatured state apparently characterized by incomplete exposure of tryptophan residues, when comparing the maximum emission fluorescence of cardosin A at pH 12 (λ_{max} = 354 nm) and with free tryptophan (about 357 nm) (Burstein *et al*., 1973). As for the acid induced state of cardosin A, incomplete exposure of tryptophan residues is also found, even though it represents a more folded state.

```
GSAVVALTNDRDTSYFGEIGIGTPPQKFTVIFDTGSSVLWVPSSKCINSK   50
ACRAHSMYESSDSSTYKENGTFGAIIYGTGSITGFFSQDSVTIGDLVVKE  100
QDFIEATDEADNVFLHRLFDGILGLSFQTISVPVWYNMLNQGLVKERRFS  150
FWLNRNVDEEEGGELVFGGLDPNHFRGDHTYVPVTYQYYWQFGIGDVLIG  200
DKSTGFCAPGCQAFADSGTSLLSGPTAIVTQINHAIGANEELQVDCNTLS  250
SMPNVSFTIGGKKFGLTPEQYILKVGKGEATQCISGFTAMDATLLGPLWI  300
LGDVFMRPYHTVFDYGNLLVGFAEAA                           326
```

Figure 4.2: Amino acid sequence of cardosin A and distribution of ionisable amino acids. Amino acids with pK_a near 8 are highlighted in red (Tyr), green (Cys) and blue (Lys). Small chain is in bold. Regions corresponding to the catalytic triads (DTG/DSG) and the RGD and KGE domains are highlighted in grey and underlined. Cardosin A protein data bank accession number is 1b5f (Berman *et al*., 2000).

To further investigate the molecular changes of cardosin A with varying pH, size exclusion chromatographic (SEC) studies were carried out. In Figure 4.1C, D and E, retention times and correspondent apparent molecular weights of cardosin A, eluted at different pH, are shown as well as the correspondent chromatographic elution profiles. Overall they indicate that from pH 5-11 a gradual increase of the protein volume is clearly noticeable as seen by the decrease in elution volumes, ranging from 25.4 ml for the native state up to 21.3 ml for cardosin A at pH 11. This would correspond to a 64.2 kDa globular protein elution volume comparing to a 38.8 kDa molecular estimated weight for native cardosin A. Nevertheless, at pH above 12 two distinct chromatographic peaks become evident, being named peak 1 and 2 (Figure 4.1E). Electrophoretic analysis of these two forms at pH 13 (Figure 4.1F) was carried out under both denaturing and non-denaturing conditions. Analysis of silver stained SDS-PAGE gel showed that peak 1 is totally constituted by the large chain of cardosin A, whereas peak 2 is constituted by the small polypeptide chain even though a faint band corresponding to the larger polypeptide chain could be detected. The estimated molecular weights at pH 13 could not be extrapolated with confidence by SEC due to the void volume overlapping with the elution volumes of cardosin A at this pH. Nevertheless, extrapolation of molecular weights from elution volumes of cardosin A at pH 12 could still be calculated indicating 89.6 and 55.3 kDa, suggesting monomer molecular associations into oligomers that imply high monomer instability in solution.

Silver staining of native gel of cardosin A at pH 13 showed several protein bands on lanes 1 and 2, corroborating monomers association in alkaline solutions (Figure 4.1F).

With all evidences shown, the alkaline induced denaturation state of cardosin A is characterized by a first transition phase where a gradual destabilization of the secondary and tertiary structure can be observed, possibly due to repulsion of negative charges, resulting in an expansion of the molecular volume. Subsequently a second transition, at pH above 10, is seen both by spectroscopic and hydrodynamic measurements where almost full dissociation of the heterodimer takes place and with further destabilization of the secondary and tertiary structure. Finally the alkaline denatured state of cardosin A is not considered a fully unfolded state.

At the same time, cardosin A retains activity in the pH range 2.5-7.5 and seems to be mainly related with protonation states of both catalytic aspartates since only slight structural effects are detected in this pH range. If it was not due to the changes in protonation of the catalytic aspartates, that at some point preclude catalysis, it would be possible to follow structural changes in cardosin A, pointing to dissociation of the chains, alongside with the functionality of the catalytic apparatus. This could shed light in subunit structure/activity relationships that in this study were barred by activity loss due to protonation effects. Concerning the monomeric pepsin it has been established in a number of studies that the alkaline-denatured state is composed of a tightly folded *C*-terminal lobe with a substantial amount of nonnative secondary and tertiary structures, and a *N*-terminal lobe with residual structure (Lin *et al*., 1993; Favilla *et al*., 1997; Kamatari *et al*., 2003). Considering the structure homology between these enzymes, the underlying issues are if analogous unfolding events could occur in heterodimeric cardosin A structure and to what extent could this unfolding events affect the functionality of the active site cleft.

A thermodynamic approach on the structural stability of cardosin A according with pH will hopefully help explain the observed pH-induced effects and discuss results with available thermodynamic data on other related APs. So far, thorough structural analysis of pH induced unfolding on APs have only been pursued for pepsin and other few pepsin-like APs, as noted in Section 1.2.6, and HIV1-protease, being the first monomeric and the latter, homodimeric APs.

4.1.1.2 Temperature induced unfolding – Thermodynamic analysis

Previously the structural stability of cardosin A has been briefly characterized at pH 5 (Oliveira, 2001). Cardosin A at pH 5 goes through thermal transition around 70 ^{o}C and analysis of the DSC scans suggested that the excess heat capacity could be deconvoluted into two components corresponding to two elementary two-state transitions. Such behaviour could be explained by independent unfolding of the two polypeptide chains of cardosin A. Furthermore, CD investigations of cardosin A at this pH revealed that after thermal denaturation the polypeptide chains of this enzyme retain most of their secondary structure motifs and hence are not completely hydrated. By then the DSC curves for cardosin A at several pH values were compared regarding their transitions temperatures (T_m) (Oliveira, 2001): cardosin A DSC scans showed that this enzyme has maximum stability at pH 5, as revealed by the higher T_m value obtained. Below and above pH 4-9 the stability decreased substantially, as seen with the lower T_m values. The least stable conditions were for pH 8 and 9, with T_m values around 50 ^{o}C (Oliveira, 2001).

In this work, the pH dependence of the structural stability of cardosin A was further examined by measuring the pH dependence of the transition temperature and by estimation of thermodynamic parameters. Figure 4.3 shows a set of the excess heat capacity functions for cardosin A (solid lines) at different pH values.

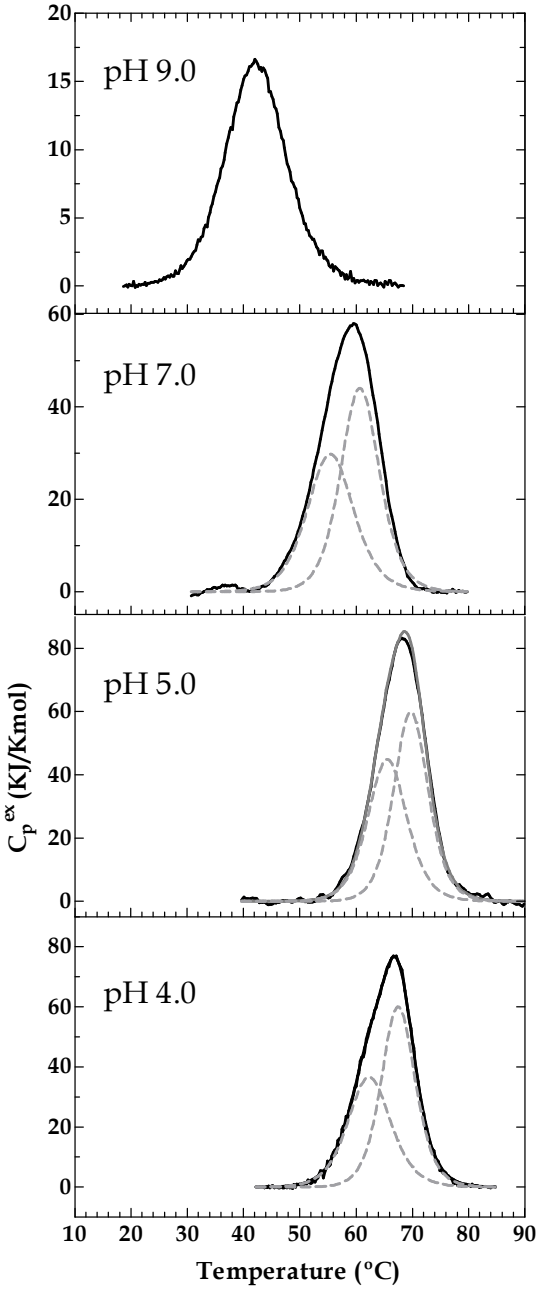

Figure 4.3: Temperature-dependence of the excess molar heat capacity of cardosin A at the different pH. Experimental traces were corrected for the chemical baseline in accordance with Takahashi *et al.,* 1981. Dashed lines represent the results of non-linear least square fittings of experimental curves to the two-state independent model as implemented in the Origin software package (Sections 3.8.1 and 3.8.3).

It is well known that a correct thermodynamic description of protein stability is possible only if its experimentally measured unfolding transition is reversible

(Privalov, 1979). For that purpose the reversibility of cardosin A thermal transitions at several pH were investigated. It was seen that in the pH range of 3-9 the extent of reversibility, measured by relative area recovery, seen on a second scan of cardosin A depended on the temperature at which the first scan was completed (before cooling the samples in preparation for the second scan). When the first scan was allowed to proceed up to a temperature at which the transition was 50 % complete, then the repeated scans showed approximately 85 % reversibility. In Figure 4.4 an example of an irreversible thermal transition effect on the chromatographic elution profile of cardosin A is given (Figure 4.4, elution 3) and compared with the elution profiles of native state enzyme (elution 1) and with the enzyme heated up to 60 ^{o}C (elution 2). It can be seen that cardosin A when heated up to 100 ^{o}C, and allowed to cool down, shows a complex chromatographic profile with several overlapped chromatographic peaks, probably a result of irreversible thermal effects such aggregation, hydrolysis of peptide bonds at Asp residues or chemical alteration of residues (Zale *et al*., 1986). Applying equilibrium thermodynamics to apparently irreversible processes has been discussed before such as for the denaturation of core protein *lac* repressor (Manly *et al*., 1985) and for the ATCase subunit (Edge *et al*., 1985). Calculations for a model system where protein undergoes a reversible denaturation followed by an irreversible step, showed that data can be analyzed according to the *van't Hoff* equation and yield parameters close to the originally assigned to the reversible system. Therefore, despite the irreversible nature of the thermal transition of cardosin A, a semi-quantitative analysis using thermodynamic models was still possible as will be demonstrated.

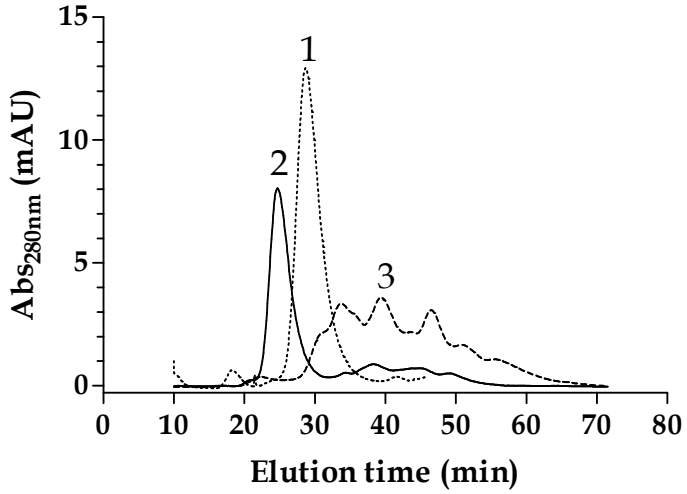

Figure 4.4: Effect of temperature in cardosin A size exclusion chromatography elution profiles. Cardosin A was analysed at pH5 at room temperature (elution 1), after heating up to 60ºC (just before T_m) (elution 2) and after heating up to 100ºC (elution 3). Chromatographies were carried out as described in Sections 3.5 and 3.5.2.

All the DSC scans of cardosin A performed in the same pH range were practically independent of the scan rate. Heating scan rates of cardosin A solutions of 28, 60 and 90 ºC/h gave similar denaturation profiles that differed in the transition temperatures by less than 0.3 ºC indicating the absence of kinetic effects under this experimental conditions. Additionally, according to equilibrium thermodynamics, any change in the oligomerization state of proteins during their thermal denaturation should produce a concentration-dependence of T_m (Fukada *et al*., 1983; Edge *et al*., 1985). In this work, the dependence of the thermal transition temperature on the protein concentration was analyzed at pH 5. It turned out that the thermal transition temperature of cardosin A does not depend on the protein concentration, differing by less than 0.4 ºC within the 0.3-4.8 mg/ml range. The constancy of T_m may have two meanings. It can be taken as evidence that cardosin A is already dissociated by the time temperature reaches T_m or that the protein remains dimeric after thermal denaturation. The hypotheses of cardosin A dissociation before reaching the T_m seems unlikely as indicated by SEC of cardosin A at pH5 and heated near the T_m (Figure 4.4, elution 2), where cardosin A is

eluted as a main single peak. This may be taken as evidence that the thermal denatured protein remains in the same oligomerization state as the native protein, i.e. thermal denaturation of cardosin A is not accompanied by simultaneous dissociation of the folded dimer to the unfolded monomers. This is considered to be an unusual situation for multimeric proteins but can be explained assuming that the thermal denaturation of cardosin A would occur without its complete unfolding and, in particular, that some structural elements responsible for the interaction between the two different cardosin A chains would remain. In a previous study native and thermally denatured cardosin A was analyzed by CD in the Far UV region and by intrinsic fluorescence, at pH 5 (as shown below in Section 4.1.1.3), confirming that a considerable portion of the secondary and tertiary structure elements still persist in the temperature induced denatured state (Oliveira, 2001).

Regarding the excess heat functions at different pH values (Figure 4.3), none of the DSC curves shown here display the typical asymmetry that would be expected if denaturation were to follow the form N$m \rightarrow m$U, where N and U represent the native and denatured protein, respectively, *i.e.* the protein remaining oligomeric up to the denaturation temperature and then undergoing simultaneous denaturation and dissociation to *m* molecules of the unfolded dimeric protein (Manly *et al*., 1985).

Comparisons of the two enthalpies – *van't Hoff* (ΔH_{VH}) and calorimetric (ΔH_{Cal}) enthalpy of unfolding, can be used with high accuracy to test the nature of the transition processes (Becktel et al., 1987). If both enthalpies are equal, the unfolding is considered to be a two-sate process, whereas if $\Delta H_{VH} < \Delta H_{Cal}$ the unfolding process is considered to be more complex, as being not monomolecular or involving intermediate states. So, to further evaluate the unfolding process for cardosin A at different pH for the obtained calorimetric profiles the ratio of the *van't Hoff* enthalpy and the calorimetric enthalpy were calculated according to the equations below (Privalov, 1979):

$$\Delta H_{VH} = \frac{4\,RT_m^2\;C_p^{\max}}{\Delta H(T_m)} \qquad (1)$$

$$\Delta H_{Cal} = \frac{\Delta H_{VH}}{\Delta H(T_m)} \quad (2)$$

where $C_p^{\max}$ refers to the maximum heat capacity upon unfolding and $\Delta H(T_m)$ to the enthalpy change at the thermal transition temperature (T_m). The *van't Hoff* enthalpy (ΔH_{VH}) is obtained indirectly from the temperature dependence of an equilibrium constant whereas the calorimetric enthalpy is obtained from a direct measurement of a heat change (ΔH_{Cal}).

Hence, for pH values of 3-8 the ratios averaged 0.45 ± 0.04 and 1.08 ± 0.05 for pH 9. The value 0.45 could stand for a two-state transition of a dimeric protein (Munson *et al*., 1996), while the value of 1.08 is close to 1, which would be found for two-sate transition of monomeric protein (Privalov *et al*., 1974). The two-chain structure of cardosin A strongly implies that this scheme reflects the independent unfolding of its two chains. Even though cardosin A polypeptide chains are of different size, they have similar relative folds as can be seen analyzing the secondary structures content of each chain, calculated from cardosin A crystallographic data (Frazão *et al*., 1999; Berman *et al*., 2000) (Table 4.2). Since the chains are of different size but with similar folds, they are expected to have different unfolding enthalpies as seen for cardosin A.

Table 4.2: Relative secondary structure content of cardosin A. The large chain represents the 31 kDa polypeptide whereas the small represents the 15 kDa peptide. Cardosin A protein data bank accession number is 1b5f (Berman *et al*., 2000).

Chain	Strand	α-helix	3-10 helix	Other	Total residues
Large	109	25	5	100	239
	45.6%	10.5%	2.1%	41.8%	100
Small	40	8	6	33	87
	46.0%	9.2%	6.9%	37.9%	100

The results of the deconvolution of the DSC profiles obtained by means of the Microcal software under the assumption of an independent two-sate model of unfolding are shown in Figure 4.3 (dashed lines) and Table 4.3. A good description of cardosin A thermal unfolding at different pH values was obtained.

The minor and major peaks were assigned to the smaller and larger chain, respectively. Their enthalpies were designated ΔH_S, ΔH_L and the transition temperatures as T_{mS}, T_{mL}, where the subscripts *S* and *L* stand for the short and long chains, respectively. As follows from an analysis of the DSC scans, the T_m for both chains of cardosin A remains essentially the same within the pH region from 4 to 6, whereas it decreases substantially at pH values below 3 and above 7. In the pH region where the protein is relatively reversible, the change in calorimetric enthalpy (ΔH_{Cal}) is linearly related to T_m, as can be seen in Figure 4.5. The linear relationship between ΔH_{Cal} and T_m (r = 0.99) in the pH range 3-9 provides the way to estimate changes in heat capacities values (ΔC_P) for separate cardosin A chains, since marked overlapping of the transitions in cardosin A precludes even rough estimates of ΔC_P for each chain from the total calorimetric scan. The ΔC_P upon unfolding determined from the slopes of the straight lines are ΔC_{PS} = 10.0 ± 0.4 KJ/^{o}C mol and ΔC_{PL} = 8.8 ± 0.4 KJ/^{o}C mol for the short and long chains.

Table 4.3: Thermodynamic parameters for the individual transitions of cardosin A obtained by differential scanning calorimetry at different pH values[a].

	First transition						Second transition					
pH	T_m (°C)	ΔH (T_m) (KJ/mol)	ΔC_p (KJ/°C mol)	ΔG^o (25°C) (KJ/mol)	T_s (°C)	ΔG^o (T_s) (KJ/mol)	T_m (°C)	ΔH (T_m) (KJ/mol)	ΔC_p (KJ/°C mol)	ΔG^o (25°C) (KJ/mol)	T_s (°C)	ΔG^o (T_s) (KJ/mol)
3.0	56.7	334.7	10.0	16.4	25.0	16.4	63.3	431.0	8.8	29.1	17.7	30.0
4.0	62.5	370.3		19.5	27.6	19.6	67.6	477.0		35.2	17.4	36.1
5.0	65.7	413.8		24.1	26.9	24.2	69.8	485.3		36.5	18.9	37.1
6.0	62.8	380.3		20.6	27.0	20.7	67.7	447.7		31.6	20.4	31.8
7.0	55.7	318.4		14.9	25.3	14.9	60.8	403.8		25.8	17.9	26.6
8.0	46.1	212.1		6.9	25.6	6.9	54.9	350.6		19.5	17.3	20.5
9.0	-	-		-	-	-	42.0	235.6		8.8	16.3	9.8

[a]T_m is defined as the temperature at the midpoint of the unfolding transition (the standard deviation is ±0.2 K); $\Delta H(T_m)$ is the calorimetric enthalpy of the unfolding transition with a standard deviation of ±5 %; ΔC_p is the difference between the heat capacities of the intact and denatured states obtained from the slope of the graph of the temperature-dependence of $\Delta H(T_m)$ by pH variation of T_m (the standard deviations are ±0.4 kJ/°C mol); the free energy changes, ΔG^o, were calculated with the *Gibbs-Helmholtz*, Equation 4; temperature of maximum stability, T_s, was calculated according to Equation 5 (described later below).

The molar heat capacity curve for cardosin A at pH 9 could only be approximated by a single elementary contour under the assumption of a two-state model of unfolding and the parameters of this transition correspond to the linear relationship between ΔH_{Cal} and T_m only for the long cardosin A chain.

So summing up, the structural stability of cardosin A has been thoroughly investigated by DSC: it was found out that even though the thermal denaturation of cardosin A is partially irreversible, valid thermodynamic data could be obtained within a wide pH region. Although cardosin A is a heterodimeric protein its thermal denaturation occurs without simultaneous dissociation of unfolded monomers. Moreover, in the 3-7 pH region the excess heat capacity could be deconvoluted into two components corresponding to two elementary two-state transitions, suggesting that the two chains of cardosin A unfold independently.

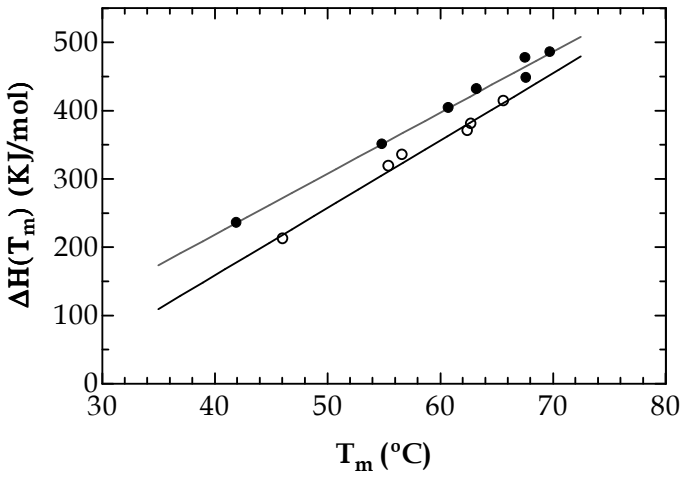

Figure 4.5: Linear least-squares determination of the apparent relative heat capacity (ΔC_p) of the denatured states of cardosin A subunits. The T_m and ΔH values for the long (solid circles) and short (open circles) chains of cardosin A were obtained by fitting the DSC traces measured at seven pH values to an independent two-state model.

In fact, SEC, intrinsic fluorescence and CD experiments described in chapter 4.1.1.1, Figure 4.1, show that between the pH range 5–11 cardosin A undergoes gradual loss of secondary and tertiary unfolding resulting in a molecular expansion volume without effective chain separation. Nevertheless, at pH values above 11, dissociation of cardosin A chains inevitably occurs, corroborating the independent unfolding of each chain. At pH 9 only a single thermal transition was observed, assigned to the long chain, implying that the short chain at this pH is already extensively denatured but at this point still not dissociated from the larger chain. Nevertheless, increasing alkaline conditions for cardosin A leads to chain dissociation, probably due to weakening of the intra subunit interactions and/or to increasing of electrostatic repulsion forces. On the other hand, cardosin A below pH 2 undergoes a steep loss of structure as indicated by spectroscopy (Figure 4.1A). Yet, due to strong protein absorption to the SEC matrix equilibrated in strong acidic solutions, no information is available concerning the acid induced denatured state in relation with subunit dissociation. However, the proposed transition model accounts for no change in the oligomerization state of acid induced denatured state of cardosin A.

4.1.1.3 Structural and thermodynamic characterization of cardosin A at pH 5

Cardosin A activity and stability was characterized at several pH values. Further studies were conducted at pH 5 were cardosin A displays near maximum activity and stability. Figure 4.6A, depicts the temperature-dependence of normalized fluorescence intensity. Taking into account the proposed denaturation mechanism following analysis of the DSC data, the first-order temperature derivatives of fluorescence data at pH 5 was analyzed to further support these evidences. For this purpose it was assumed a superposition of two independent transitions, from each of which a two-sate reversible equilibrium process between native and denatured states follows. The best global fit of these data (Figure 4.6A, solid line) was achieved accepting that the parameter of these assumed contours satisfies the following equation:

$$\Delta H(T_m) = \frac{4\,RT_m^2}{\Delta T} \qquad (3)$$

where ΔT is the width at the half-height of the contour (Privalov *et al*., 1974) and with values taken from the analysis of calorimetric profiles (Figure 4.3). The means of the enthalpy of the denaturation and transition temperatures for the short and the long chain of cardosin A are $\Delta H_S = 377.4 \pm 17.2$ kJ/mol, $\Delta H_L = 496.2 \pm 25.1$ kJ/mol, $T_{ms} = 66.2 \pm 2.1$ ºC and $T_{mL} = 69.6 \pm 1.5$ ºC, agreeing well with the calorimetric results (Table 4.3).

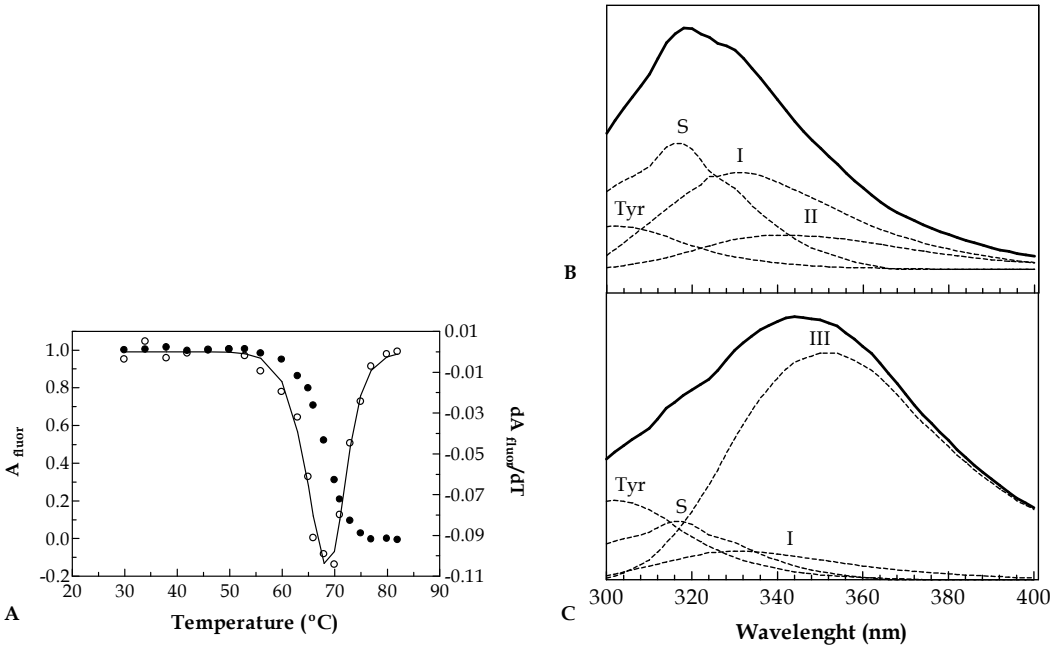

Figure 4.6: Characterization of the thermal denaturation process and of the denatured state of cardosin A at pH 5 monitored by intrinsic fluorescence: (A) represents the thermal denaturation monitored by measuring the normalised area under the fluorescence spectrum recorded at a heating rate of 30 ºC/h (closed circles). The open symbols are first-order temperature derivatives of the corresponding experimental data and the solid line is the result of fitting the data assuming a two-sate unfolding model (Section 3.6.2); (B and C) represents the fitting of the experimental fluorescence spectra of intact (B) and thermally denatured (at 85 ºC) (C) cardosin A (symbols) to the theoretical model of discrete states of tryptophan residues in proteins (solid lines). The fitting represents the sum of the spectral components Tyr, S, I, II and III (dashed lines) (Section 3.6.6).

Figure 4.6B and C, solid lines, shows the emission spectra of intact and thermally denatured (up to 85 ^{o}C) cardosin A at pH 5, excited at 295 nm. Overall, comparison of two spectra shows that thermal denaturation of the protein results in a red shift of the tryptophan fluorescence spectrum maximum by more than 20 nm, from 323 to 347 (closer to the spectrum position of free tryptophan in water - about 357 nm), meaning that polar tryptophan environment becomes more mobile upon thermal unfolding. This usually occurs when tryptophan residues move from the hydrophobic interior to the protein surface in contact with free water molecules.

In Figure 4.6B, dotted lines, analysis of the spectrum of native cardosin A in terms of the model of discrete states of tryptophan residues (Burstein et al., 1973; Burstein, 1983; Reshetnyak et al., 2001) is shown. It can be seen that the tryptophan residues of the S and I forms provide the main contributions to the emission, as can be seen in Table 4.4. Both forms represent internal tryptophan residues forming two different types of complexes with polar groups. On the other hand analysis of the spectrum of thermally denatured cardosin A showed a pronounced red shift of the tryptophan fluorescence spectrum (Figure 4.6C, dotted line). Tryptophan residues of the form III provide the main contribution to the emission in this state (Table 4.4). This corresponds to external tryptophans residues in contact with free water molecules. As mentioned before, thermally induced conformational state of cardosin A is not a completely denatured state, and as can be seen by analysis of the deconvoluted spectrum, part of the tryptophan residues are located inside the remaining protein structure represented by the forms S and I.

Table 4.4: Characterization of cardosin A native and thermally denatured states. Spectral components relative abundance of tryptophan fluorescence emission in cardosin A was determined according to the theoretical model of discrete states of tryptophan residues in proteins (Section 3.6.2 and 3.6.6).

Cardosin A state	λ_{max} (nm)	Spectral components (%)				
		Tyr	S	I	II	III
Native	319.7	11.5	34.5	38.4	15.6	-
Thermally denatured	349.5	13.3	10	7.1	-	69.6

The CD spectra of intact and thermally denatured cardosin A at pH 5 was already obtained (Oliveira, 2001) and is shown in Figure 4.7B. It was then described that intact cardosin A displays a broad minimum at 217 nm, which is characteristic of β-structure proteins, as is the case for cardosin A and pepsin-like APs. Profound changes in the CD spectra take place after the thermal denaturation of cardosin A, clearly indicating a change in structure: the broad band centred at 217 nm decreases in intensity and changes in shape while the positive band at 197 nm is transformed into a negative band at 200 nm. It was then reported that thermal denaturation leads to almost complete disappearance of β-turns and an increase in unordered structure, while the contents in the number of α-helical and β-strand structures remain unchanged. The data together with intrinsic fluorescence spectroscopy results confirm the perception that cardosin A does not undergo complete unfolding upon thermal denaturation.

As performed with fluorescence spectroscopy, the process of thermal denaturation of cardosin A at pH 5 was monitored by changes in molar ellipticity at 215 nm, since at this wavelength both the pre-transitional and pos-transitional baselines show only a weak temperature-dependence. Figure 4.7A, closed symbols, shows the normalized thermal denaturation curve of cardosin A at pH 5. Reversibility of the thermal denaturation was verified previously (Oliveira, 2001), and revealed to be similar with the results obtained in the calorimetric investigations, demonstrating approximately 80 % reversibility. The first-

order temperature derivative of the unfolding curve (open symbols in Figure 4.7A) was analyzed assuming that it consists of two components (solid line), in the same way as in the case of the analysis of the fluorescence data. The means for the enthalpy of denaturation and transition temperatures for the short and the long chains of cardosin A obtained from the best global fit of this curve (solid line) are ΔH_S = 391.2 ± 15.9 kJ/mol, ΔH_L = 493.7 ± 16.7 kJ/mol, T_{ms} = 66.7 ± 1.4 °C and T_{mL} = 71.2 ± 1.3 °C, agreeing well with both the calorimetric and fluorescence results as can be seen altogether in Table 4.5.

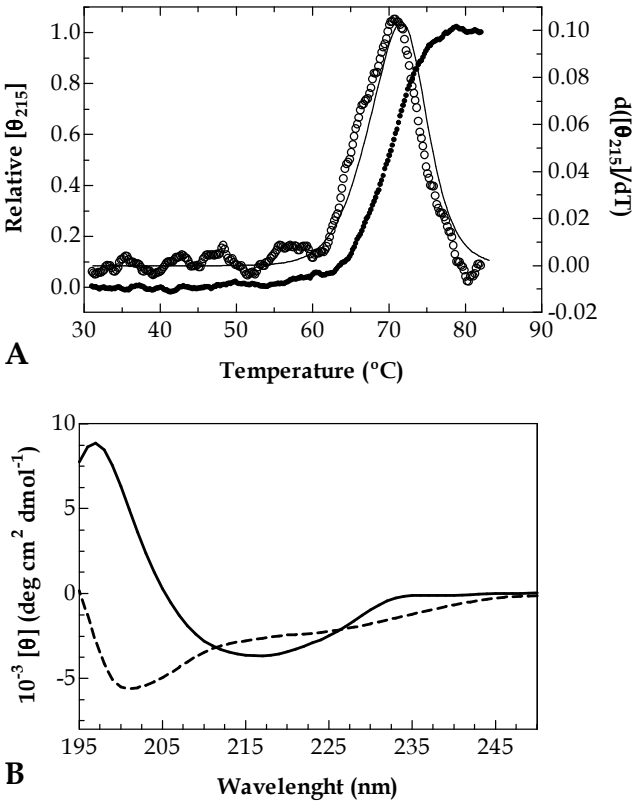

Figure 4.7: Characterization of the thermal denaturation process and of the thermal denatured state of cardosin A at pH 5 by CD: (A) Closed symbols represent the normalized changes measured at 215 nm. The open symbols are first-order temperature derivatives of the corresponding experimental data and the solid line is the result of fitting the data assuming a two-sate unfolding model (Sections 3.7.2 and 3.7.5); (B) represents the far ultraviolet CD of intact cardosin A at 25 °C (solid line) and thermally denatured at 85 °C (dashed line) at pH5 (adapted from Oliveira, 2001).

Table 4.5: Estimates of ΔH and T_m values for the short and long polypeptide chains of cardosin A.

	Short Chain		Long Chain	
	ΔH_S (kJ/mol)	T_{ms} (°C)	ΔH_L (kJ/mol)	T_{mL} (°C)
DSC	413.8	65.7	485.3	69.8
Fluorescence	377.4 ± 17.2	66.2 ± 2.1	496.2 ± 25.1	69.6 ± 1.5
CD	391.2 ± 15.9	66.7 ± 1.4	493.7 ± 16.7	71.2 ± 1.3

Overall, the characterization of the thermal denaturated state of cardosin A at pH measured both by CD and fluorescence showed that cardosin a does not undergo complete unfolding upon heat denaturation. There are still secondary structure elements in the thermal denatured sate as well as buried tryptophan residues in the protein structure. The proposed thermal denaturation mechanism dictates that this heterodimeric enzyme follows an independent two-sate model of unfolding, which implies the independent unfolding of its two polypeptide chains. Nevertheless, during heating the occurrence of molecular aggregation events still remains to be investigated. To investigate this, scattered light intensity of cardosin A was carried out and results presented in Figure 4.8, closed symbols clearly show that the process of fluorescence intensity change with temperature is accompanied by an increase in the aggregation of cardosin A molecules at temperatures above 69.9 °C, near the melting temperature found initially for cardosin A at pH 5, as reflected in the increase in the scattered light intensity.

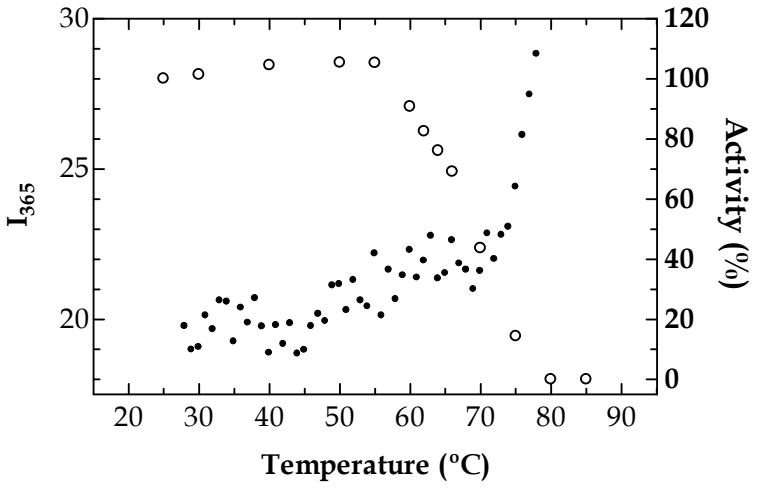

Figure 4.8: Thermal denaturation of cardosin A at pH 5 monitored by activity and by intensity of light scattering: Closed circles (left axis) represent the relative intensity of scattered light at 365 nm (Section 3.6.2); Open circles (right axis) represent the inactivation of cardosin A recorded at a heating rate of 1 °C/min (adapted from Oliveira, 2001).

As for cardosin A activity, in Figure 4.8, it can be seen that full activity is maintained up until 55 °C, before aggregation events occur and just before changes in fluorescence and ellipticity are detected, around 60 °C (Figure 4.6 and 4.7, respectively). After that, activity is gradually lost apparently accompanying fluorescence emission and ellipticity changes. Finally, at 80 °C cardosin A is completely inactivated.

So, overall, cardosin A activity during heating seems to be affected by both changes in secondary structure content as well as tertiary structure rearrangements of the molecule that gradually and in parallel occur during heating. It was also shown that as these changes were detected, both polypeptide chains are unfolding independently without chain separation and that the small polypeptide unfolds more rapidly than the longer one. From available cardosin A structure information (Frazão et al., 1999) it is known that the catalytic triads and neighbour residues are located in the long polypeptide chain (Figure 1.4). There are no clear evidences in data obtained with pH and thermal induced

unfolding that correlate the proteolytic activity of cardosin A with both chain structure integrities.

The experimental parameters characterizing the denaturation of cardosin A can be used to calculate the free energy change for denaturation at any temperature, $\Delta G^o (T)$, using the modified *Gibbs-Helmholtz* equation (Becktel *et al*., 1987) :

$$\Delta G^\circ (T) = \Delta H (T_m) \left(1 - \frac{T}{T_m}\right) + \Delta C_p \left[(T - T_m) - T \ln\left(\frac{T}{T_m}\right)\right] \quad (4)$$

and the temperature where ΔG^o (T) reaches a maximum, T_s, using:

$$T_s = T_m \exp\left[-\frac{\Delta H (T_m)}{T_m \Delta C_p}\right] \quad (5)$$

The protein stability curves for both cardosin A chains calculated with these equations and the experimental parameters obtained in this work are shown in Figure 4.9. The values of T_m, ΔG^o at T_s and at 25 ^{o}C are given in Table 4.3. The calculated values of T_s testify to the relative hydrophobicity of cardosin A, assuming that this enzyme can, in principle, be cold-denatured (Privalov, 1990). The stability of cardosin A chains at 25 ^{o}C is 21-38 kJ/mol. This seems to be reasonable because it is known that in biological processes energy is exchanged, as a rule, by portions of 17-34 kJ/mol (Privalov, 1979).

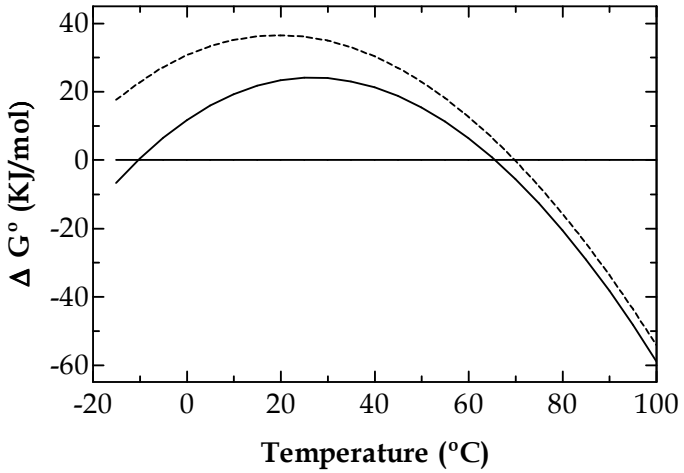

Figure 4.96: Temperature-dependence of the Gibbs energy for the short chain (solid line) and the long chain (dashed line) of cardosin A at pH 5.0. Curves were plotted according to Equation 4 and 5.

4.1.2 General discussion

Cardosin A behaviour induced by changes in pH was initiated in a previous investigation, and in this work, additional data was presented that allowed a thorough discussion in this issue. Overall, cardosin A is active in the pH range of 2.5-7.5, with maximum activity at 4.5 presenting the typical bell shaped curve suggesting the summation of the curves for the two catalytic aspartate titrations. The combination of spectroscopic techniques allowed following structure changes and hence the characterization of the acid and alkaline induced conformational states for cardosin A.

In the pH range where cardosin A was active, no conformational changes were detected (fluorescence and CD studies) and the activity seemed to be controlled by the protonated state of the catalytic aspartates.

There are many consistent data that shows how the structure of the active site can be adapted for the function in such a wide range of pH. In general it is believed that the optimum pH results from two main factors – organization of the titrable groups with the

three dimensional structure and amino acid composition (Andreeva *et al*., 2001; Alexov *et al*., 2004).

In fact, and according with the accepted mechanism of pepsin-like enzyme function, Asp215 has to be charged, whereas Asp32 has to be protonated (Davies, 1990). A remarkable property of this catalytic center is adaptation for functioning in a wide range of pH, from pH 1 up to pH 7. Accordingly, different pH optima are found in APs. Most eukaryotic APs have optimal activity in the pH range 2-4, while others have slight different preferences. For example, retroviral proteases catalyze best in the range of 5.5-7.5, whereas cardosin A near pH 5 (Table 4.1). These slight pH preferences are associated with their action in diverse environments where a physiological adaptation of the same structural motif was necessary. This probably has been the main reason for the evolutionary success of this class of proteases. Pepsin, acting in food digestion of mammal stomach, has optimal pH around 2 that allows operating in its natural acidic environment, HIV-1 proteinase functioning in cytosol, functions best at pH 7 whereas the cardosin A works best near pH 5, suitable for vacuolar environments. In APs the adaptations to different environments result from the assistance of amino acid residues adjacent to the active site that have to preserve the charged state of Asp215 and the protonated state of Asp32 (Andreeva *et al*., 2001). Retroviral proteases, on the other hand do not possess such regulating systems, due to its homodimeric nature. In fact, it can be considered a remarkable imitation of mammalian aspartic proteases. HIV-1 protease is a homodimer - more genetically economical for the virus with its two-fold symmetric active site with the two Asp residues - one with high pK_a and other with low pK_a, as seen in Table 4.1. As for cardosin A and pepsin, heterodimeric and monomeric proteins, pK_a values are slightly different.

Also, a correlation has been found from numerical calculations that correlate the optimum pH and acid/base ratio. Proteins having acidic optimum pHs have a tendency for a acid/base ratio greater than one (Alexov, 2004), providing favourable electrostatic environments for acids. In Table 4.1, the acid/basic ratios are listed for cardosin A, pepsin and HIV-1 proteinase. It can be seen that the ratios demonstrate this correlation for cardosin A and other APs, where more acidic enzymes present the highest acid/base ratios.

Cardosin A, below and above the pH range 2.5, and 7.5, becomes inactive and suffers major conformational changes. It was demonstrated that the acid and alkaline induced unfolding for cardosin A are different, being the first monophasic and the second biphasic. Likewise, the alkaline and acid induced states for cardosin A are different. The alkaline induced state is characterized by chain dissociation as well as incomplete unfolding, whereas in the acid induced state, a more compact molecule is presented and where dissociation is not expected. To evaluate the role of each chain in unfolding, further studies were performed and the thermodynamic stability of cardosin A determined. It was shown that its thermal denaturation occurs without simultaneous dissociation into monomers and that in the 3-7 pH range the two polypeptide chains unfold independently. Additionally, the small chain was shown to be less stable than the large chain.

At pH 9, DSC studies showed no thermal transition for the small chain, attesting that in these conditions, it has already suffered massive unfolding but still associated with the large chain, as predicted by SEC experiments. As pH rises, the non cooperative independent unfolding of the chains initiates with weakening of intrachain interactions and strengthening of charge repulsions, leading to inevitable dissociation.

The independent unfolding of cardosin A chains can be compared with previous unfolding studies with some APs. Regarding pepsin, a single chain enzyme, alkaline inactivation is due to a selective denaturation of its *N*-terminal lobe, and there is evidence for higher *C*-terminal stability comparing with the *N*-terminal portion of pepsin. In the alkaline induced state, pepsin has a folded *C*-terminal and an unstructured *N*-terminal portion of the molecule (Privalov *et al*., 1981; Lin *et al*., 1993). On the other hand HIV-1 protease, a homodimeric chain protein, the structure-based thermodynamic analysis predicts the existence of regions of the protease with only marginal stability and a high propensity to undergo independent unfolding. In particular the flaps region was predicted to be easily affected by relatively small perturbations (Todd *et al*., 1998). This information provided new insights for the development of therapeutic agents that inactivate HIV-1 protease by competing with natural substrates or by destabilizing the structure to the point in which the enzyme loses its activity.

Also, this independent unfolding has already shown to play a physiological role in pepsin. It was hypothesized that taking advantage of the pH gradients inside the crypts of the stomach and of the transit of pepsin from the parietal cells, pH induced

intermediate sates in pepsin participate in transportation to the stomach lumen and allow for pepsin activity control. In HIV-1 protease no physiological inferences have been made, since apparently no such reports in literature are found. This can be explained because HIV-1 structural and functional studies performed up until now were focused mainly in the development of HIV-1 inhibitors for pharmaceutical purposes. Nevertheless regarding HIV-1 protease, its low pH tolerance and its homodimeric nature, arising by duplication of the same gene, all point to an enzyme programmed to have specific physiological role in a specific environment. This protease performs the postranslacional processing of viral poliprotein precursors to the active viral enzymes and structural proteins in infected human cells. All this results, and as said before, to be more genetically economical for the virus.

Cardosin A physiological role has yet to be totally clarified as discussed before (Section 1.2.7.2, Introduction) but recently new supplementary information was added concerning the interaction of phospholipase *Da C2* domain with cardosin A RGD and KGE-mediated binding (Simões *et al*., 2005). This suggested concerted and/or synergistic actions in degenerative processes such as in stress responses, plant senescence and/or pollen-pistil interactions. It is possible that the heterodimeric nature of cardosin A, the different chains stabilities and the mentioned pH tolerance range are the result of an AP adaptation to multiple physiological functions in its specific environment. The partial unfolding of the molecule could facilitate the *C2* domain binding process, and at the same time, promote or prevent catalysis, since the catalytic aspartates are located in the long chain. Results shown up until now, nevertheless do not shed light on the regulatory role of the small chain in cardosin A activity and stability.

4.2 Structural and thermodynamic study of cardosin A – The effect of acetonitrile

4.2.1 Dissection of acetonitrile effects on cardosin A

As was described in the previous chapter, the effects of pH and temperature in cardosin A were investigated. A structural and functional characterization was performed and results compared with data available for APs like pepsin and HIV-1 proteinase. The investigation concluded that cardosin A structure when destabilized suffers independent unfolding of its polypeptide chains, with the long chain being more stable than the short chain. Since in this AP the catalytic aspartates are both located in the long polypeptide chain (Figure 1.4, Introduction), an investigation concerning the role of the short chain in maintaining the active site cleft functional seemed pertinent. Nevertheless, for pH induced effects, such study could not be fully performed due to the strict dependence of cardosin A activity, as well as for the remaining APs, with the protonation state of catalytic aspartates precluding structure and function correlations. Furthermore the study of the temperature effects in cardosin A also precluded proper structure and functions correlations due to protein aggregation and other thermal irreversible effects. In this sense, for structure and function studies of cardosin A, another destabilizing agent was needed, where conformational induced changes events could be trailed with changes in activity.

Mixtures of water and acetonitrile (CH_3CN) have been extensively studied as of purely scientific interest and also because of its importance to applied chemistry (Guillaume *et al*., 1997; Carrea *et al*., 2000).

Acetonitrile is considered by some authors to be one of the "typically non-aqueous" solvents. This is due to the presence of a partial negative charge on the nitrogen side, which enables hydrogen bonding with water through nitrogen (Bertie *et al*., 1997; Fadnavis *et al*., 2005), determining the solubility of water in acetonitrile. The hydrogen bonding in the mixture destroys the three-dimensional hydrogen bond network of water, even though the hydrogen bonds between water and acetonitrile are considerably weaker.

A great deal of information about its structure and properties has already been obtained and agreed in that any water-acetonitrile solution is not a simple homogenous mixture of two components, being possible to distinguish three different main acetonitrile concentration regions (Davis *et al*., 1987; Marcus *et al*., 1991).

Based on excess volumes, viscosity, dielectric constants and acid–base properties the presence of three structurally different regions over the water/acetonitrile mixture composition range was inferred. The bounds for such regions, in mole fractions of acetonitrile, were $0 \leq \chi \leq 0.2$ (up to 40 % acetonitrile, v/v), $0.2 \leq \chi \leq 0.75$ (from 40 % to about 90 %) and $0.75 \leq \chi \leq 1.0$ (90 % up to 100 %).

In the first region, on the water rich side, the water structure remains more or less intact as acetonitrile molecules are added interstitially into cavities in this structure, the so called homogenous system. At 40 % acetonitrile up to 90 %, the so called microheterogeneity sets in, meaning that relatively large water clusters are formed and, therefore, water molecules tend to be mostly surrounded by water, while acetonitrile molecules attach to acetonitrile. This tendency originates from the stronger hydrogen bonding between water molecules than with acetonitrile. For 90 % acetonitrile concentrations and higher, individual water molecules interact with individual acetonitrile molecules.

Cardosin A has already been partially characterized concerning the effects of 10 % acetonitrile in cardosin A storage and operational stabilities (Sarmento, 2002; Oliveira, 2001; Sarmento *et al*., submitted). It was seen that the addition of 10 % acetonitrile does not induce specificity alterations in cardosin A and does not change the pH activity dependence, ruling out that solvent molecules would change the ionization state at the active site residues (Sarmento *et al*., 2004b).

Initially the studies of cardosin A in organic solvents where carried out having in mind practical application. Later, a thorough study of acetonitrile induced effects in cardosin A structure and activity seemed appealing to pursue structure and function studies. In this sense the study of the effects of growing concentrations of this destabilizing agent were carried out envisaging that the solvent perturbation would result in partial disruption of secondary and tertiary structures that could be correlated with changes in catalytic activity.

4.2.1.1 Acetonitrile induced unfolding- structural analysis

To follow conformational changes occurring through the addition of acetonitrile, fluorescence and CD studies were recorded.

The fluorescence of proteins is a very sensitive indicator of the microenvironment of the tryptophan residues and so it is a valuable tool to explore tertiary structure changes in proteins. The parameter more accurately affected by environmental factors is the wavelength of maximum emission (λ_{max}). In general, a red shift of emission wavelength of an enzyme is an indication of an increase in the polarity of the microenvironment of the tryptophan residues. It is also well known that both the intensity and wavelength of the maximum emission (λ_{max}) of tryptophan fluorescence depend on solvent polarity. When studying fluorescence at different solvent compositions these effects have to be taken into account and minimized. This can be done by comparing enzyme fluorescence with the fluorescence of model tryptophan derivatives such as N-acetil-L-tryptophan ethyl ester (ATEE) (Kijima *et al*., 1996), as described in Equation 6, below:

$$\Delta \lambda_{em} = \lambda^{ATEE} - \lambda^{Cardo\,\sin A} \qquad (6)$$

Where λ^{ATEE} and $\lambda^{Cardo\,\sin A}$ represent the emission wavelength of ATEE and cardosin A, respectively and $\Delta \lambda_{max}$ represents the change in emission wavelength.

Figure 4.10 summarizes the fluorescence emission data for cardosin A and ATEE in acetonitrile. It can be seen that the λ_{max} for ATEE does change, or more specifically, it decreases steadily with increasing acetonitrile content.

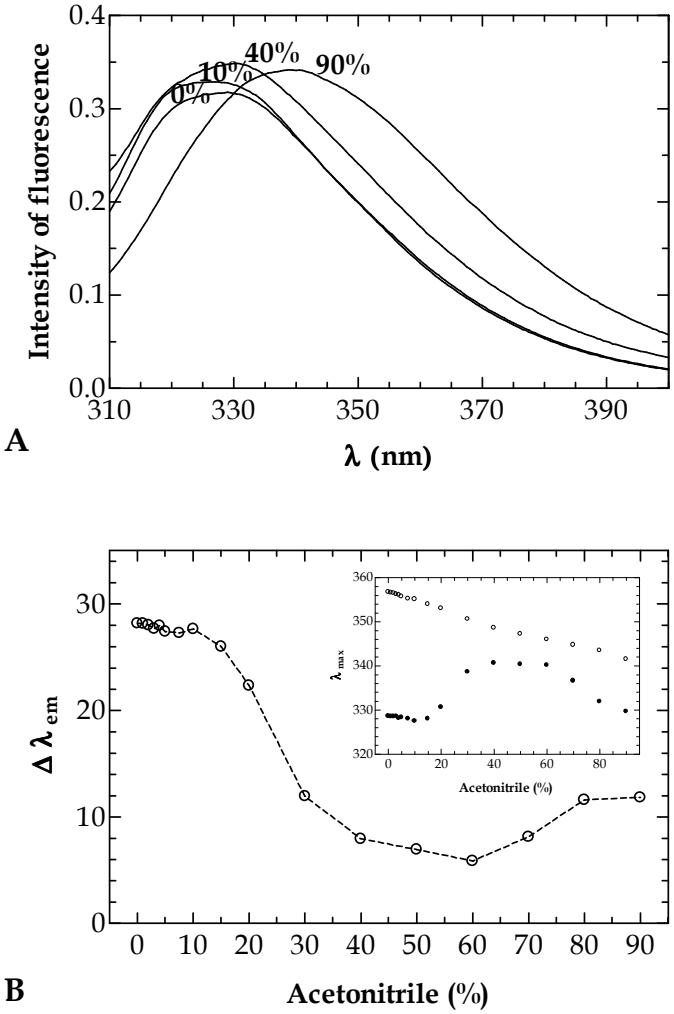

Figure 4.10: Cardosin A acetonitrile induced effects monitored by intrinsic fluorescence. (A) Cardosin A intrinsic fluorescence at some acetonitrile concentrations after 1h incubation at 25 ºC. (B) Δλmax of cardosin A in the presence of growing concentrations of acetonitrile, where Δλmax = λATEEmax - λEmax, to minimize the direct effects of solvent polarity on the tryptophan emission. λATEEmax and λEmax represents wavelength maximum of ATEE and cardosin A, respectively. The inset shows wavelength of the emission maximum of ATEE (open circles) and cardosin A (closed circles) in the presence of acetonitrile (Section 3.6.3).

As for cardosin A, the addition of increasing concentrations of acetonitrile does not change the emission maximum in a continuous manner. It can be seen that mild concentrations of the organic solvent, up to 10 %, do not change the tryptophans environment in cardosin A, presenting spectra identical do native cardosin A. Above 15 % acetonitrile, a pronounced change in emission fluorescence occurs up to 40 % acetonitrile corresponding to exposure of cardosin A tryptophan residues to polar environment. In the acetonitrile concentration range of 40-60 % tryptophans exposure to aqueous medium is less abrupt ending in maximum tryptophan exposure, even though tryptophan residues exposure is not complete. This can be clearly noticed in Figure 4.10B, where at 60 % acetonitrile $\Delta\lambda_{max}$ is low, but still different from zero. When $\Delta\lambda_{max}$ is close to zero the environmental polarity of tryptophan residues is considered to be similar to that of ATEE, indicating that the tryptophan residues are fully exposed to the solvent, as a result of unfolding of the peptide chains. Finally, above 60 % acetonitrile concentrations a blue-shift in λ_{max} is clearly observed, as indicated by the increase in $\Delta\lambda_{max}$, indicating a rearrangement of tryptophan residues in cardosin A towards less polar environments. As suggested by different emission fluorescence at 90 % acetonitrile, cardosin A at lower water contents suffers a non native rearrangement of the tryptophan residues in the molecule towards less polar environment.

Concerning the dependence of cardosin A concentration in acetonitrile induced effects, the fluorescence spectra of different cardosin A concentrations in some acetonitrile concentrations were collected. As can be seen in Figure 4.11, no important differences were found for acetonitrile induced effects at two different cardosin A concentrations in what exposure of tryptophan residues was concerned.

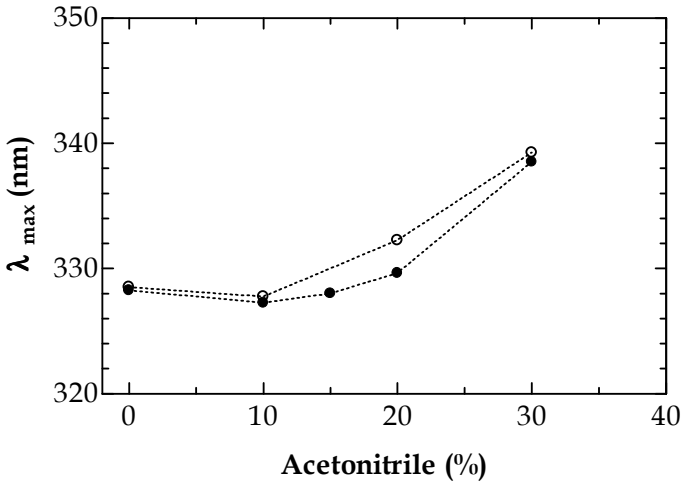

Figure 4.11: Effect of different cardosin A concentrations in the wavelength of the emission maximum of cardosin A incubated at some acetonitrile concentrations. Open circles stand for cardosin A at 11.9 x 10-4 mM and closed circles at 4.77 x 10-4 mM (Section 3.6.3).

A further examination of the effects of acetonitrile on the enzyme was carried out by CD measurements to follow the structural changes in the secondary structures. In the far-UV region, CD spectra of proteins are particularly sensitive to protein secondary structure. The Figures 4.12A and B show the CD spectra in the far-UV region and CD signals observed at 195 nm for cardosin A, respectively, and measurements were made in 0-60 % aqueous acetonitrile range. At concentrations of acetonitrile above 70 % cardosin A precipitates at concentrations of 0.1 mg/ml or above precluding reliable spectroscopic measurements. The use of lower enzyme concentrations to minimize precipitation adversely affected the reliability of the spectra.

In the acetonitrile concentrations range studied, profound changes in the CD spectra of cardosin A take place. Spectra of intact cardosin A, has shown before (Oliveira, 2001), is characteristic from all-β proteins, even though these spectra are more dissimilar and weaker than those from helical proteins (Greenfiled *et al*., 1969). Nevertheless, the general characteristics, and shared with intact cardosin A spectra, are a negative band about 216 nm and a positive band of comparable magnitude near 195 nm. It is possible to see (Figure 4.12A) two differences in acetonitrile induced changes followed by CD. For

spectra obtained with higher acetonitrile concentrations (20 and 60 % acetonitrile), there is a gradual decrease of the positive band near 195 nm and a decrease of the negative band around 216 nm indicating progressive loss of secondary structure of cardosin A. On the other hand, the 5 % acetonitrile spectra show an increase, even though small, of the negative band around 216 nm and maintaining a strong positive band at 195 nm. This seems to indicate that cardosin A in this acetonitrile concentration, has a slight increase in secondary structure content.

It is known that there is a strong difference near 195 nm between the CD of folded and unfolded proteins, providing, in principle, an ideal probe for monitoring protein unfolding transitions. Nevertheless the use of this wavelength range is usually precluded by strong denaturants absorption below 200 nm. However this is not the case with acetonitrile enabling conformational changes monitorization at this wavelength. Therefore, it can be seen in Figure 4.12B, by ellipticity recorded at 195 nm (closed circles), such as in the range 1-10 %, that there is a small increase in structure content of cardosin A. In fact, there have been some similar reports of solvent induced structure and there is not really a good understanding of this phenomenon. It seems that for many proteins, low concentrations of denaturant can promote folding and, considering that the folded structure is a very subtle balance between the forces within a protein and those between a protein and the solvent, this could be possible (Ulijn *et al*., 2002; Luo *et al*., 1997).

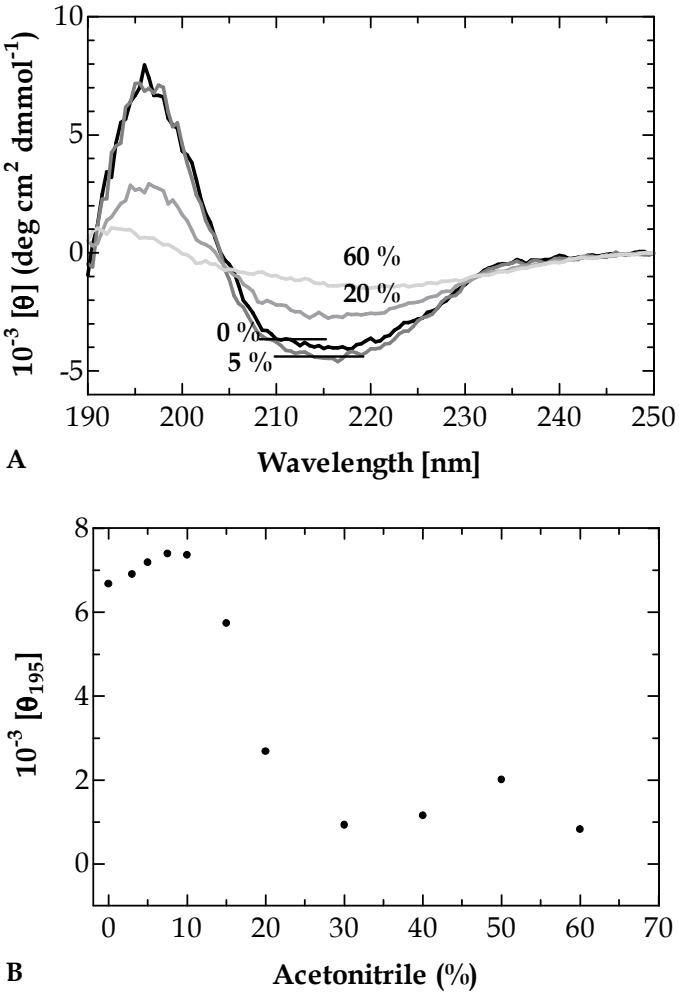

Figure 4.12: Cardosin A acetonitrile induced effects monitored by CD. (A) Cardosin A far-UV CD spectra at various acetonitrile concentrations after 1h incubation at 25ºC. (B) CD signals observed at 195 nm (Closed circles) in the presence of different concentrations of acetonitrile (Sections 3.7.3).

However, between acetonitrile concentrations above 15 %, a large gradual decrease in ellipticity at 196 nm to almost complete loss of signal was observed. This suggests that treatment of cardosin A with higher acetonitrile concentrations results in near loss of secondary structure content. The signals observed above 50 % acetonitrile most probably lack real meaning, since in these acetonitrile concentrations the changes can be affected by small errors in cardosin A concentration determination.

The behaviour of cardosin A upon gel filtration was analyzed under both native and denatured conditions to yield information concerning protein volume and to test for possible acetonitrile induced dissociation events. Figure 4.13 shows cardosin A elution profiles at different acetonitrile concentrations. Native cardosin A elutes as a single peak with an elution volume (V_e) of 25.5 ml. When the acetonitrile concentration is increased, up until 20 % acetonitrile, no shift in the V_e is observed. All the changes in tertiary and secondary structure detected by fluorescence and circular dichroism induced by acetonitrile do not alter cardosin A hydrodynamic volume. This means that rearrangement of tryptophan residues in cardosin A and changes in secondary structure content occur without transition to a more compact or expanded intermediate. At 50% acetonitrile, cardosin A elutes also as a single peak, indicating no dissociation of the 31 and 15 kDa polypeptide chains. This peak is shifted to a greater V_e of 37.2 ml, indicating a more compacted molecule in this state. Fluorescence data indicate that at 50 % concentration of acetonitrile tryptophans in cardosin A shifted to more polar environments in the molecule, but according to gel filtration, this rearrangement of the molecule presumably is followed by tightening of the molecule. Nevertheless, fluorescence data above 50 % is not consistent, due to lack of peak symmetry and also due to an increase of interactions of cardosin A with the column matrix, hampering a clear analysis at higher acetonitrile concentrations.

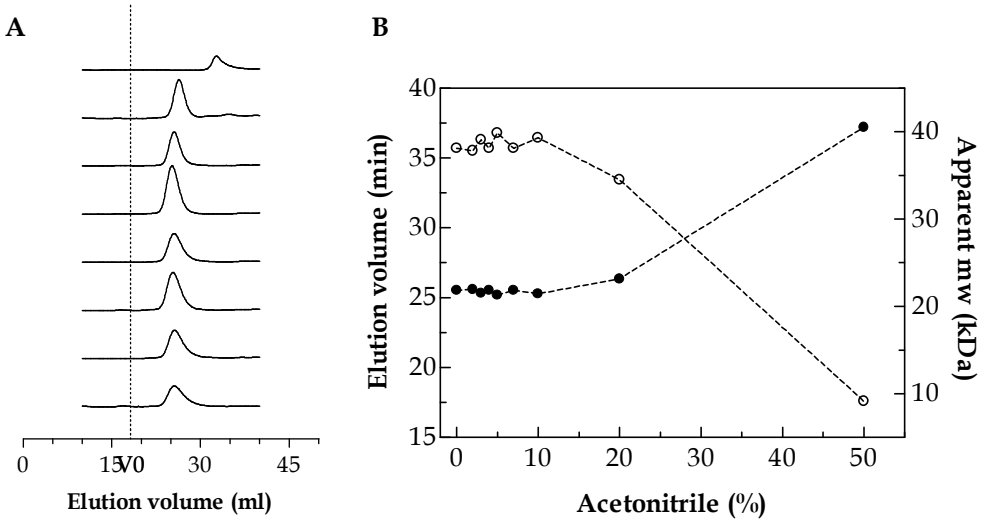

Figure 4.13: Size-exclusion chromatographic elution profiles of cardosin A at different acetonitrile concentrations: (A) Elution profiles of cardosin A at different acetonitrile concentrations. From bottom to top acetonitrile concentrations are 0, 2, 3, 4, 7, 10, 20 and 50 % (Section 3.53). V0 indicates de column void volume; (B) Elution volumes and estimated molecular weight values of cardosin A as a function of acetonitrile concentration (Section 3.5.5).

Overall, the study of acetonitrile induced effects in cardosin A structure were investigated by intrinsic protein fluorescence, to monitor for changes in tertiary structure, CD, for monitorization of secondary structure changes and SEC to detect heterodimer dissociation and molecular volume changes. The changes in signal (ellipticity in CD and wavelength of maximum emission in fluorescence) show that the profiles are similar, at least up to 40 % acetonitrile, suggesting no intermediaries, detectable or stabilized, in the acetonitrile induced unfolding transition. At low acetonitrile concentrations (below 10 %) only small secondary structure changes were found in the whole cardosin A molecule and above this range, up to 40 %, extensive unfolding of the enzyme is observed with no chains dissociation. Higher acetonitrile concentrations seem to induce in cardosin A a nonnative-like reorganization of the molecule. These effects in cardosin A have already been described for other proteins in aqueous-organic mixtures and in pure organic solvents (Griebenow et al., 1996). It can be concluded that there are two effects simultaneously at play: as the organic solvent content in the medium is raised, the tendency of a protein, cardosin A in this investigation,

to denature increases. And on the other hand, as the water content in the medium declines, the protein conformational mobility diminishes. This causes that at high acetonitrile concentrations, although the high propensity to denature, the capacity to undergo denaturation is impaired. For cardosin A aggregation events occurring at higher acetonitrile concentrations impaired proper structure analysis. For some other enzymes it was possible to observe that they are more catalytically active in pure organic solvents than in aqueous-organic mixtures (Klibanov, 1997).

To know whether cardosin A activity performance in acetonitrile follows the described trend and to better characterize the acetonitrile induced conformational states, the catalytic activity in the presence of acetonitrile was investigated.

4.2.1.2 Acetonitrile induced unfolding – Activity studies

Usually, addition of small amounts of a water-miscible solvent has little effect on the biocatalyst activity and stability. In some cases, modest concentrations of this solvent show an enhanced enzyme activity and stability (Butker, 1979; Vazquez-Duhalt *et al*., 1983; Batra *et al*., 1994). When the concentration is increased, most water-miscible solvents have an inhibitory effect on the biocatalyst. However, in the presence of pure organic solvents, striking activity can sometimes be observed (Bromberg *et al*., 1995; Castro, 1999). To compare the structural data of cardosin A in acetonitrile with the effects in the active site ability to hydrolyze peptide bonds, activity of cardosin A in the presence of acetonitrile was investigated.

4.2.1.2.1 Inactivation and reversibility experiments

Cardosin A activity in the presence of growing concentrations of acetonitrile was investigated (Section 4.2). Figure 4.14A, open circles, shows the results obtained for cardosin A activity in the presence of increasing amounts of acetonitrile. Hydrolysis rate of Lys-Pro-Ala-Glu-Phe-Phe(NO2)-Ala-Leu peptide was measured. As can be seen at least up to 10 % acetonitrile concentration, an important increase of the catalytic activity was observed. This remarkable increase of catalytic activity (up to 182 % for 5 % acetonitrile) can not be explained by changes in tertiary structure of cardosin A that could in some way be beneficial for the catalytic function, since no differences in spectra were found (Figure 4.10). In contrast, a small increase in secondary structure content was observed (Figure 4.12B). These subtle structural changes could be responsible for the increased cardosin A activity, but other factors could also be at play and will be discussed below.

Above 10 % acetonitrile concentration, the catalytic activity decreased dramatically up to about 30 % acetonitrile, where complete inactivation was observed. This dramatic decrease of cardosin A activity was accompanied by abrupt changes in tertiary and secondary structures. For low water medium, no activity was detected as has been reported for some enzymes (Luo et al., 1997). As for cardosin A activity determination in pure acetonitrile, experiments were impossible to carry out due protein aggregation.

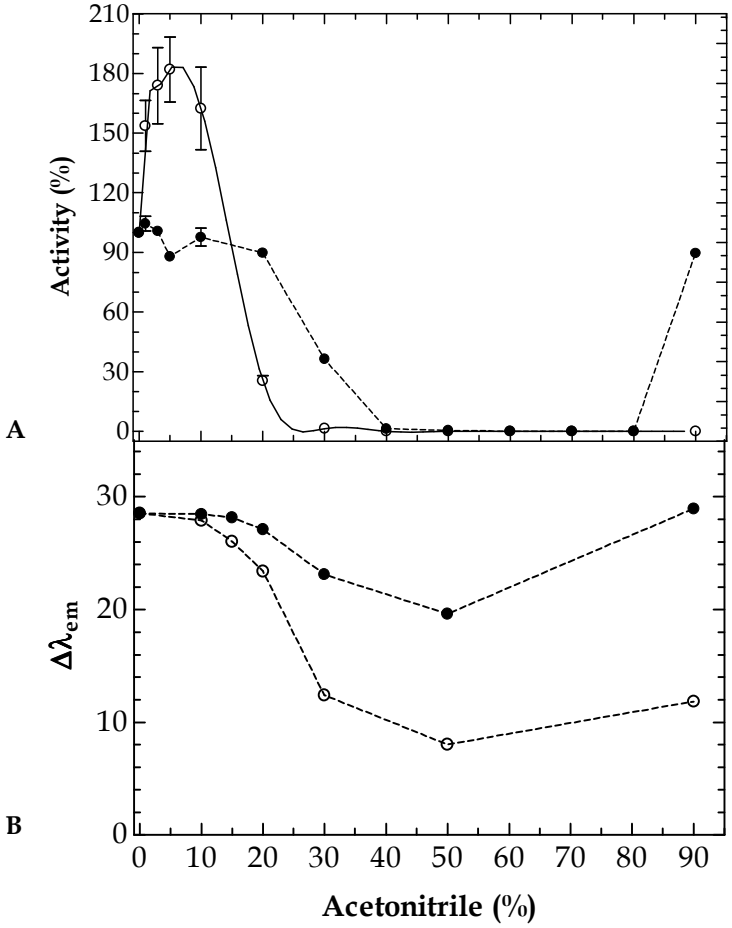

Figure 4.14: Cardosin A acetonitrile induced effects monitored by activity and reversibility followed by activity and fluorescence experiments. (A) initial rates of hydrolysis of Lys-Pro-Ala-Glu-Phe-Phe(NO2)-Ala-Leu catalysed by cardosin A in the presence of varying acetonitrile concentrations. Residual activity was measured after 1h incubation at 25 ^{o}C (open circles) and also after dilution of the acetonitrile in cardosin A incubation solution - 200 fold (closed circles) (Section 3.4.2). (B) changes in emission wavelength for cardosin A incubated at different acetonitrile concentrations measured in the presence of acetonitrile (open circles) and after dilution (up to 200 fold) with buffer (closed circles) (Section 3.6.4).

To investigate the nature of acetonitrile effects, reversibility was inspected. In Figure 4.14A, closed circles, it can be seen that cardosin A activity in up until 20 % acetonitrile is almost fully recovered, after dilution of cardosin A solution. Additions of aqueous buffer to the reaction mixture containing inactivated cardosin A by acetonitrile concentrations above 40 %, resulted in absence of activity recovery. Nevertheless, dilution of inactive

cardosin A solution at 90 % resulted in recovery of activity. At this concentration, cardosin A is probably trapped in a more rigid conformation enabling it from undergoing to an irreversibly unfolded state enabling that when the organic solvent is washed out probably results in recovery of native-like active conformation. To confirm this, reversibility tests of acetonitrile induced effects in cardosin A structure were carried out by intrinsic fluorescence. Figure 4.14B, closed circles, shows the changes in wavelength of maximum fluorescence emission of cardosin A after dilution from acetonitrile solutions. It can be seen the reversibility of the acetonitrile induced effects, measured by fluorescence, follows the trend for reversibility of the catalytic activity, demonstrating the close relationship between the tertiary fold and the integrity of the active site cleft. Moreover after dilution from a 90 % acetonitrile solution, where recovery of activity was seen, the wavelength of emission maxima of cardosin A matches the native conditions. This supports the reversibility of acetonitrile activity inhibition seen for cardosin A at 90 %, suggesting that at lower water content this enzyme is trapped in a more rigid conformation that prevents irreversible unfolding, on contrary to what was observed for example at 40 and 50 % acetonitrile.

The dependence of cardosin A concentration in acetonitrile induced effects in the catalytic activity was also investigated. Figure 4.15 shows the results obtained herein and discussed below.

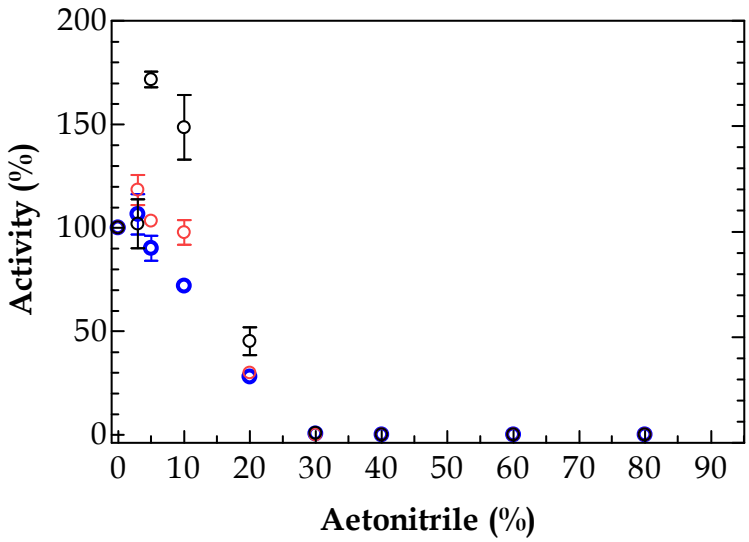

Figure 4.15: Effect in activity of cardosin A incubation in acetonitrile at different protein concentrations Residual activity was measured after 1h incubation at 25 °C with the following cardosin A concentrations 34.7 x 10-4 mM (blue open circles), 6.95 x 10-4 mM (red open circles) and 1.39 x 10-4 mM (black open circles) (Section 3.4.2).

It can be seen that complete inactivation occurs at the same acetonitrile concentration (30 %) for the differently concentrated cardosin A solutions. This could mean that for higher acetonitrile concentrations the drastic structure effects that inactivate the enzyme are not shielded by higher number of protein /protein contacts that occur in more concentrated protein solutions. Furthermore, up to 10 %, the enhancement of activity, discussed before, is more noticeable for lower cardosin A concentrations. Similarly, suggesting a direct effect in cardosin A molecules that is more facilitated in more dilute protein solutions.

4.2.1.2.2 Kinetics of cardosin A hydrolysis in acetonitrile and substrate studies

The kinetics of cardosin A in aqueous-organic media were determined from the initial hydrolysis velocities of Lys-Pro-Ala-Glu-Phe-Phe(NO_2)-Ala-Leu substrate by HPLC analysis (Section 3.4, Chapter 3). Since calculation of kinetic constants of an enzymatic reaction in organic solvents requires the knowledge of the functional active-site concentration in organic solvents, cardosin A concentration in the acetonitrile range

studied was determined by active site titration using pepstatin, a tight-binding APs inhibitor (Knight, 1995).

Table 4.6 shows the fitted kinetic parameters for this substrate at 25 ºC and in different acetonitrile concentrations. For cardosin A in aqueous buffer, the determined K_m is 0.090 mM, falling in the range with previously published kinetic studies for this enzyme, 0.04 mM (Sarmento *et al*., 2003) and 0.108 mM (Veríssimo *et al*., 1996). Results show that the K_m values for cardosin A slowly increased as the low solvent concentrations were elevated (up to 10 % acetonitrile) revealing that in this range the substrate binds more weakly to the enzyme active site.

At this point, as showed in Figure 4.14A, open circles, cardosin A exhibits significant activation, as represented by the increase in maximum velocities (V_{max}) and turnover numbers (K_{cat}). This activation can not be explained by the drop in cardosin A/substrate affinity. The lower affinity can be the result of subtle conformational changes at the active site, induced by acetonitrile or the result of acetonitrile induced structure or solvation changes in the substrate. On the other hand, these subtle structural changes in the enzyme or in the substrate could somehow favour the catalytic efficiency.

As for higher solvent concentrations, such as for 20 % acetonitrile, where there is a steep drop in activity, K_m value increases deeply, about 54 fold. V_{max} and K_{cat} estimates for these conditions could not be estimated with high accuracy as can be seen with the estimated errors. In this sense, no conclusion can be drawn with safety concerning the maximum velocities registered and the number of substrate molecules converted into products per unit of time. Nevertheless, spectroscopy studies at this concentration of acetonitrile showed that cardosin A exhibits a dramatic conformational change, comparing with 0 % acetonitrile, intuitively explaining the deficient catalytic performance, specifically the lack of affinity between substrate and enzyme active site cleft.

Table 4.6: Kinetics of Lys-Pro-Ala-Glu-Phe-Phe(NO2)-Ala-Leu hydrolysis by cardosin A in selected concentrations of acetonitrile (Section 3.4.2.4).

Acetonitrile (%)	K_m (mM)	V_{max} (mM.min^{-1})	K_{cat} (min^{-1})	K_{cat}/K_m (mM^{-1}.min^{-1})
0	0.09003 ± 0.00655	0,00441 ± 0.00009	356,0 ± 7,5	3954,0
1	0.12 000± 0.009693	0.00646 ± 0.00014	nd	*nd*
3	0.18670 ± 0.01325	0.00827 ± 0.00020	nd	*nd*
5	0.18740 ± 0.02044	0.01035 ± 0.00038	508,1 ± 18,5	2711,3
10	0.30920 ± 0.03343	0.01056 ± 0.00047	470,0 ± 20,8	1519,9
20	4.92000 ± 0,03343	0,01776 ± 0,00524	816,5 ± 240,5	166,0

nd: not determined

The parameter that truly represents the effect of an organic solvent on an enzyme is the specificity constant, K_{cat}/K_m, being considered a catalytic efficiency parameter. As can be seen in Table 4.6 the catalytic efficiency of cardosin A in acetonitrile is always lower than in aqueous conditions, even though in the same range for 5 and 10 % acetonitrile higher activity is seen. Further, at 20 %, the catalytic efficiency is highly diminished, about 20 fold, corroborating with the high K_m value and the deficient activity, mostly caused by important structure changes. However, for 5 and 10 % acetonitrile, the K_{cat}/K_m indicates lower catalytic efficiency parameters, but in the same range as found for cardosin A without solvent addition, apparently contradicting the high catalytic activity observed (higher than in aqueous conditions). At the same time, K_m values at these concentrations indicate less substrate affinity for cardosin A active site binding.

The observed activity and the determined kinetic parameters seem to point three main factors acting on cardosin A: higher substrate solvation, and therefore less accessibility to the enzyme; some acetonitrile molecules interaction with the active site cleft resulting in the optimization of the biocatalyst performance, albeit lower affinity for the substrate, explaining the observed K_{cat}/K_m values; and likewise, subtle structure changes in cardosin A molecule away from the active site that could also favour catalysis.

4.2.1.2.3 Substrate induced changes with acetonitrile

It was discussed that the enzyme/substrate binding can be affected by structure changes induced by acetonitrile in the substrate, Lys-Pro-Ala-Glu-Phe-Phe(NO_2)-Ala-Leu. This substrate is synthesized and based on a chromophoric hexapeptide selected for being related with the primary structure of bovine *k*-casein that flanks the hydrolyzed peptide bond by cardosin A. The peptide bond is Phe105-Met106 and known to initiate the milk clotting process in milk (Raymond *et al*., 1973).

Small peptides often present secondary structure elements, even though tertiary structure is not expected. To investigate the structure of this peptide, CD experiments of the peptide were carried out in aqueous conditions and in the presence of some acetonitrile concentrations.

Concerning the CD spectrum of the peptide in aqueous conditions presented in Figure 4.16, it is possible to see a pronounced minimum between 195 nm and 200 nm, and week signals above 210 nm. These features point to unordered peptide, that is to say, these signals are typical for peptides with little, or none, ordered secondary structure (Schmid, 2005). Additionally, CD spectra of the synthetic peptide with different acetonitrile concentrations are very similar to the aqueous spectrum. They present the same minimum between 195 nm and 200 nm and even weaker signals above 210 nm. This suggests that the synthesized octapeptide in aqueous solution has no or few ordered structure, and that addition of acetonitrile does not considerably alter the peptide.

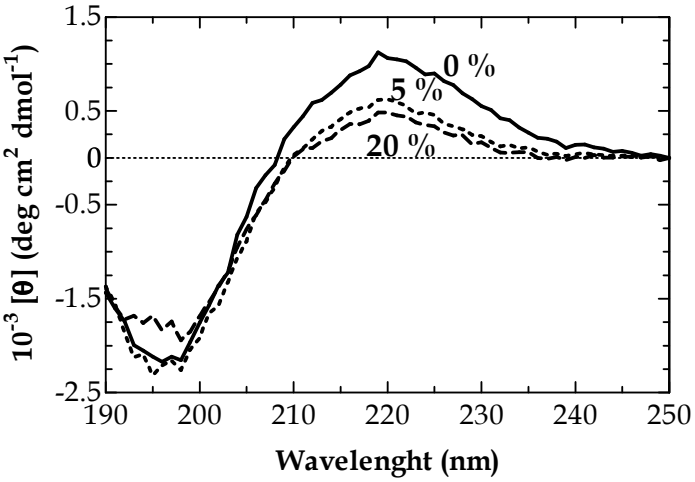

Figure 4.16: Circular dichroism spectra of Lys-Pro-Ala-Glu-Phe-Phe(NO2)-Ala-Leu, pH 5.0 in 0, 5 and 20 % acetonitrile solutions. Experiments were carried out as described in Section 3.7.3.

These results suggest that there are no acetonitrile substrate induced changes, therefore no alteration in the substrate structure occur that could be responsible for the increased catalytic efficiency of cardosin A.

Furthermore, substrate solvation of protein and peptides has been investigated and results show that there is no single parameter that can serve a universal and reliable predictor of the solvent's protein-dissolving ability (Chin *et al*., 1994; Bell *et al*., 1995). Usually the *Michaelis Menten* constant, K_m, is used to estimate the relative affinity of the substrate for the enzyme active site compared with solvent molecules. In this sense results obtained for cardosin A suggest that increasing concentrations of acetonitrile increase the substrate solvation, that is to say, increase the affinity for the solvent. Some preliminary solubility experiments were carried out concerning the solubility of the peptide and point to an increase of peptide solubility with the presence of acetonitrile, substantiating with high peptide solvation. When the substrate presents higher affinity to the solvent, weaker affinity to the enzyme is expected, and consequently the catalytic performance is negatively affected. Nevertheless, for low acetonitrile concentrations (1-10 %), the increase in the catalytic performance cannot be solely explained by differences in substrate solvation, neither by substrate structure alterations.

4.2.1.3 General discussion

In this section an investigation the effects of growing concentrations of acetonitrile in cardosin A was presented.

Structural studies in the range of acetonitrile concentrations up to 10 % showed that no substantial changes in the structure were detected, even though a slight induce of secondary structure content was seen. At the same time, gel filtration showed that cardosin A global conformation is conserved, suggesting that the subtle changes in secondary structure content could have a role in the higher activity seen, perhaps favouring the flexibility of the active site and consequently, the formation of ES complex.

Enhancement of activity caused by addition of low quantities of organic solvent has also been observed for other enzymes, but no general explanation has been accepted. However, some hypothesis can be proposed: occurrence of minor changes in structure or flexibility in cardosin A that could help catalysis, and also substrate solvation effects could be applied here. If the substrate would be less solvated by certain acetonitrile concentrations, the activity would be expected to be higher, less driving force for association with the enzyme (Yennawar *et al*., 1994). If this was true, K_m values were expected to be lower, comparing to strictly aqueous conditions, but this was not the case for cardosin A since K_m values in the range of 1-10 % acetonitrile were shown to be slightly higher, revealing a more weakly enzyme/substrate binding compared with solvent molecules (Table 4.6). This would imply that cardosin A activity could fall down. In this sense, substrate solvation effects are not expected to be main responsible for the striking cardosin A activation at low acetonitrile concentrations, since clearly other factors have major role in the activity seen.

Most kinetic studies have shown that the behaviour of enzymes in non aqueous media follow conventional models. However, the values of kinetic parameters usually are very different from those for the same enzyme and substrate in aqueous media. There appears to be a general tendency for K_m values to increase with water concentration in a given solvent (van Erp *et al*., 1991). For cardosin A, at acetonitrile concentrations up to 10 % the opposite trend is verified.

Additionally, k_{cat}/k_m showed that in the low acetonitrile concentration range, the catalytic efficiency is diminished, contrasting again with the high activity seen. If substrate solvation effects are ruled out, the striking activation can be associated with the subtle secondary structure adjustments in cardosin A. This can suggest long range structural changes effects and/or direct effects at the active site geometry. For cardosin A peptide substrate, Lys-Pro-Ala-Glu-Phe-Phe(NO_2)-Ala-Leu, CD studies revealed that in aqueous conditions this peptide has unordered structure and that addition of the organic solvent does not alter the peptide structure. This rules out any effect that the alterations in the substrate could have in enhancing the catalytic performance. On the other hand, K_m estimates suggest an increase in substrate solvation that would imply a decrease in catalytic performance, and not, as seen, an important activation of the enzyme. All the evidences suggest, therefore, that at mild acetonitrile concentrations, subtle structure adjustments occur that directly induce better catalytic performances. At the same time, results shown in Figure 4.15 suggest a direct acetonitrile effect probably in the active site region. When cardosin A is more concentrated the active site seems to be shielded from acetonitrile induced effects on activity, since the activity is not enhanced. The possible binding of acetonitrile molecules to the active site region could also explain why kinetic parameter estimates failed to explain acetonitrile effects in activity.

In fact, the catalytic apparatus of APs is characterized for allowing extensive substrate/active site interactions, as discussed before. In a previous work (Frazão *et al*., 1999) the substrate binding pockets of cardosin have been identified by analogy with renin (for which structure of enzyme-inhibitor complex had already been determined). Docking results suggested 127 atomic contacts including five putative hydrogen bonds involving altogether 30 cardosin A residues that presented close contacts (within 4 Å) with the scissile peptide bond. This clearly confirmed the complexity of the catalytic apparatus interactions in cardosin A, such as occurs in the other pepsin-like APs. This complexity can be the reason for the failure in clearly explaining the acetonitrile induced effects in activity.

It can be assumed that the main reason for the enhancement of activity on cardosin A in the acetonitrile concentration range of 0-10 % is a direct effect in the active site. Indeed previous reports on acetonitrile effects in other enzymes (none belonging to the APs) already showed that binding of acetonitrile molecules to active sites and other regions of

the molecule could occur (Fadnavis *et al*., 2005; Fitzpatrick *et al*., 1993). Acetonitrile is a polar, amphiphilic molecule and binding to protein implies complementary binding sites, mimicking biologic membranes. The acetonitrile binding in the enzyme active center suggests that this cleft has amphiphilic character, consistent with the preference of cardosin A to hydrolyse hydrophobic residues. It is also possible that other cavities at the surface of the enzyme could also interact with acetonitrile. In addition these interactions are not unexpected since acetonitrile molecules in this concentration range were predicted to fill up cavities in proteins, as discussed before.

Assuming direct acetonitrile effect in the active site cleft and considering the role of the flexible *flap*, from which the highly conserved residues Tyr75 and Thr77 were shown to interact and to have an important role in the hydrolysis of the substrate (Kempner, 1993; Okoniewska *et al*., 1999; Frazão *et al*., 1999; Sielecki *et al*., 1990), it would not be surprising if the flap was directly involved in the activity enhancement seen with acetonitrile.

Above 10 % acetonitrile a drop in activity was detected, up to 30 %, and seen to be irreversibly lost, and finally, for higher concentrations of the organic solvents complete inactivation was seen as well as reversibility of acetonitrile induced effects. This was accompanied by an abrupt change in the spectroscopic properties of dissolved enzyme. Fluorescence studies point to an abrupt, but gradual, displacement of tryptophan residues toward more polar environments, but still incomplete, in what concerns complete exposure to aqueous-organic media. In fact, complete unfolding is ruled out, since no important changes in the hydrodynamic molecular volume are seen. The fact that the drop in enzymatic activity occurs at the same concentration as the spectral perturbation, implies that it is the denaturation that causes the threshold inactivation of dissolved enzymes. This is illustrated by the increased *Michaelis Menten* constant K_m for the hydrolytic activity of cardosin A, that is, the enzyme shows poor affinity for binding with the substrate. Up to this point, reactivation analysis shows that cardosin A is irreversibly denatured. It is known that at increasing concentrations of the organic solvent in aqueous solution more and more water molecules from the protein hydration shell are stripped off the protein surface, until a certain critical amount of removed water molecules is reached. This follows from, in this acetonitrile concentration range, the presence of acetonitrile clusters near cardosin A molecules that will affect the protein hydrogen bonding network.

This will have implications in the side chain hydrogen bonding potential and the polarization of the backbone permanent dipole, such modifications inducing in turn an increase in the hydrogen bonding character of neighbouring polar residues (Rupley *et al*., 1991). Usually, it has been generally accepted that as the water content in the medium declines, the protein conformational mobility diminishes. As noted, for 90 % acetonitrile concentrations and higher, individual water molecules apparently interact with individual acetonitrile molecules. In this situation, the polar residues at the surface tend to construct more hydrogen bonds with each other. Some water molecules, acting as lubricants are lost or replaced by organic solvent molecules. This ultimately causes the enzyme to be more rigid, as proved by crystal structures of other proteins prepared in the presence of organic solvents, like acetonitrile (Fitzpatrick *et al*., 1993; Yennawar *et al*., 1994). This also explains why some enzymes have been reported to be more catalytically active in pure organic solvents than in aqueous-organic mixtures, revealing native like fold (Zaks *et al*., 1988; Griebenow *et al*., 1996). This is not the case for cardosin A that looses the catalytic function above 30 % acetonitrile.

Results shown here reveal that at high concentrations of acetonitrile, above 60%, cardosin A remains irreversibly inactivated. Fluorescence and CD results reveal that some packing of the molecule occur, indicated by blue shift in tryptophan fluorescence emission as well as some induction of secondary structure seen by CD. Nevertheless, the rearrangement of the macromolecular structure of cardosin A is far from native as well as the catalytic function that is lost.

Cardosin A behaviour in acetonitrile could be due to its heterodimeric nature. The fact that no dissociation is detected, at least at solvent concentrations as high as 50 %, suggests that hydrophobic interactions and the hydrogen bonding in cardosin A subunits interface are strong enough to resist to drastic changes occurring at the enzyme surface. Otherwise it could be expected that as the water content declined, hydrophobic interactions would diminish together with hydrogen bonding pattern, and eventually result in subunit dissociation. As in neat organic solvents, stability and activity of enzymes in water-organic mixtures depend not only on the properties and concentration of the organic solvent, but also on the nature of the enzyme, since enzymes from different sources can display different behaviours in the same organic solvent (Ogino *et al*., 1998; Gupta *et al*., 1997). Even though several structure/function studies in non aqueous media have been

carried out, they usual have focused on small monomeric globular proteins. Cardosin A structure/function investigations will fill a gap among the data available so far. The glycoprotein nature, where higher stabilization is expected, as seen with other glycoproteins (Zhu *et al*., 2001), the heterodimeric nature, with intrasubunits forces at play, makes cardosin A an attractive model protein for such studies.

4.2.2 Cardosin A in 10% acetonitrile

The dissection of acetonitrile induced effects in cardosin A was carried out and a global connection with the structural and functional features was possible. In cardosin A, growing concentrations of cardosin A induce at a first stage (below 10 % acetonitrile) mild secondary structure adjustments in cardosin A that seem to be directly related with the striking activation in this range. Higher acetonitrile content was seen to induce a steep unfolding of the molecule that directly impaired activity. All these evidences, nevertheless, only gave a global picture of unfolding events and also on conformational states of cardosin A. Additionally thorough structure/function correlations of enzymes performance in the presence of different concentrations of the organic solvent were made difficult by substrate solvation events, which had to be taken into account. This limitation could be avoided if studies were performed at the same organic solvent concentration, that is to say, at the same substrate solvation conditions. On the other hand, study of the time-dependent structural stability and activity transformations of proteins at a certain organic solvent concentration could add important information concerning structure and functional relationships in cardosin A.

While the kinetic aspects of the interaction of proteins and organic solvents in systems with minimal amounts of water have attracted much attention (Sakurai *et al*., 1988; Parida *et al*., 1991; Affleck *et al*., 1992; Griebenow *et al*., 1996), studies of the kinetic effect in systems with low quantities of organic solvent have been rather limited. Study of the time-dependent structural transformations of proteins in such systems is imperative not only for gaining a better understanding of the mechanism of action of organic solvent on

macromolecules, but also for the development of new assays for the manipulation and application of biocatalysts.

In this sense, special attention will be given toward the kinetics of the changes of the structural stability and enzymatic activity of cardosin A during its interaction with acetonitrile/water mixture. The 10 % acetonitrile concentration for this investigation was chosen. This acetonitrile concentration, corresponding to 0.037 molar fraction is in the range $0<x_{water}<0.1$ defined by some authors, constituting a homogenic system that facilitates results discussion. On the other hand, this acetonitrile concentration has already been tested for cardosin A concerning the operational activity stability kinetics. By then, cardosin A at 10 % acetonitrile and at 25 ^{o}C showed good operational stability, retaining most of the activity up to 700 h incubation time and, at the same time, the existence of an intermediate unfolded state for cardosin A was proposed (Sarmento, 2002; Sarmento *et al*., submited). Furthermore, this low acetonitrile concentration was shown in the previous section to induce mild structure effects and at the same time induced a better activity performance when compared with aqueous conditions. It seems reasonable that at higher acetonitrile contents, kinetic structural stability studies could be impaired by, fast unfolding, aggregation and activity loss events. Also, reversibility of acetonitrile induced effects will be lost, hampering correct thermodynamic analysis. Finally, and having in mind the different chain stabilities (discussed in Section 4.1.1.3), important information may be revealed concerning intra and interchains interactions in a multimeric AP, like cardosin A.

4.2.2.1 Structure function analysis

Fluorescence spectroscopy was used to characterize how incubation with acetonitrile affects the tertiary structure of cardosin A. It was already seen that this heterodimer molecule contains 5 tryptophan residues, one in the small chain and four in the major chain (Frazão *et al*., 1999) enabling its intrinsic fluorescence to be monitored during the evolution of the protein structure. First, possible changes in the fluorescence spectrum of

free tryptophan in 10 % acetonitrile, when incubated for several hours, were ruled out since the tryptophan fluorescence did not differ from that observed in aqueous solution.

In Figure 4.17 A and B, the emission spectra of cardosin A incubated with 10% acetonitrile for 2 min and 170 h, respectively are shown. The incubation of cardosin A through time, up to 210 h in acetonitrile induced a soft and gradual red shift of the maximum of the cardosin A fluorescence spectrum. The final value was much lower than that expected if all the tryptophan residues of the protein had become exposed to the solvent (about 357 nm). This seems to indicate that incubation in 10 % of acetonitrile does not affect dramatically the structure of cardosin A, but does involve important structural modifications inside the protein, which lead some of the five tryptophan residues present in cardosin A to move from the hydrophobic interior of the molecule to more solvent-exposed sites.

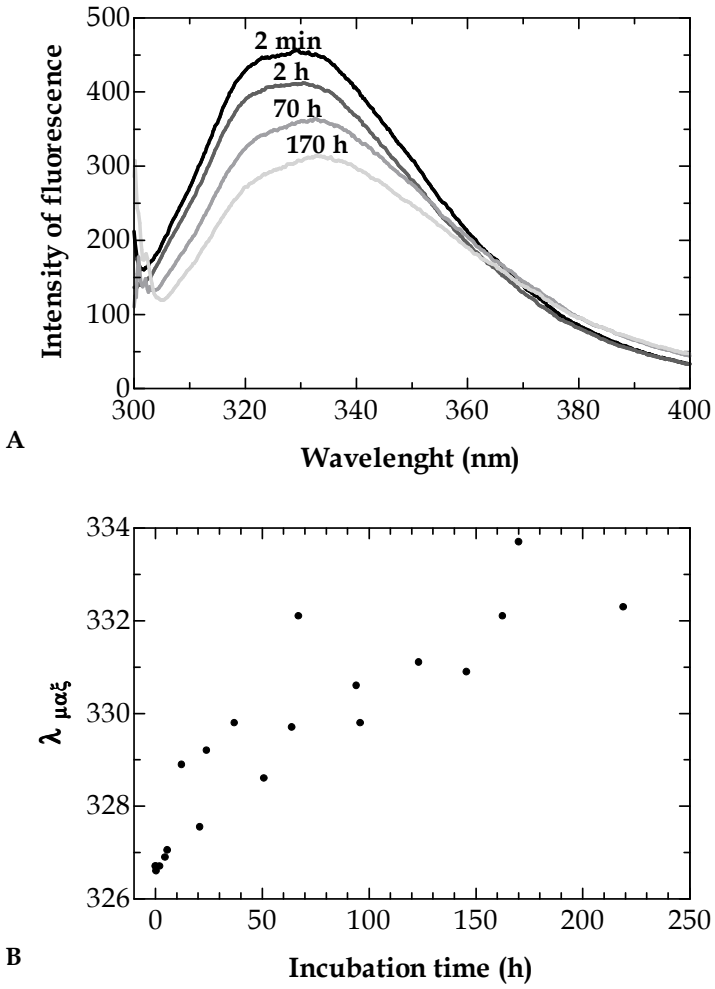

Figure 4.17: Intrinsic fluorescence monitoring of the conformational changes of cardosin A during incubation with 10 % acetonitrile, at 25ºC. (A) Cardosin A fluorescence spectra after some incubation times (Section 3.6.5). (B) Wavelength of the fluorescence emission maximum of cardosin A at various incubation times.

These structure rearrangements, with tryptophan residues shown to become more solvent exposed, could occur with the readjustment of the protein molecule volume, or even result in cardosin A dissociation, even though no chain separation was detected for 1 h incubation at the same acetonitrile content and in the whole concentration tested.

To know if the detected structure alterations induced by acetonitrile correspond to important changes in the molecular volume and also in the association state of cardosin A, SEC were carried out.

Figure 4.18 shows some chromatograms of cardosin A in 10 % acetonitrile at various incubation times. It can be clearly seen that incubation of cardosin A in 10% acetonitrile does not induce important changes in the hydrodynamic molecular volume, indicating no dissociation events and no molecular volume alterations.

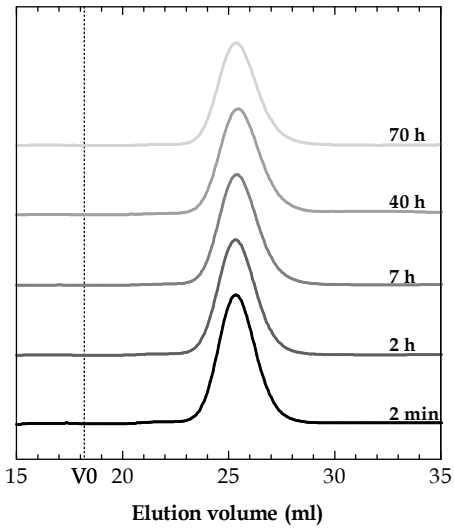

Figure 4.18: Size-exclusion chromatographic elution profiles of cardosin A at 10 % acetonitrile and with different incubation times. Experiments were carried out according to Sections 3.5.4 and 3.5.5. V0 indicates de column void volume.

Analysis of cardosin A spectra in terms of the model of discrete states of tryptophan residues in proteins (Burstein *et al.*, 1973; Burstein, 1983; Reshetnyak *et al.*, 2001) was carried out. In Figure 4.19A, it can be seen that for cardosin A in 10 % acetonitrile for 3 min, the tryptophan residues in the S and I forms (internal tryptophan residues forming two different types of complexes with polar groups) provide the main contributions to the emission. The longer incubation time (170 h) of the protein with 10% acetonitrile resulted in a soft red shift of the tryptophan fluorescence spectrum (Figure 4.19B). This shift can be modelled if only one fifth of the tryptophan residues from form I turn into form III

(external tryptophan residues in contact with free water molecules). Thus, with evidences shown it is possible that only one of the five tryptophans of cardosin A changes its environment dramatically.

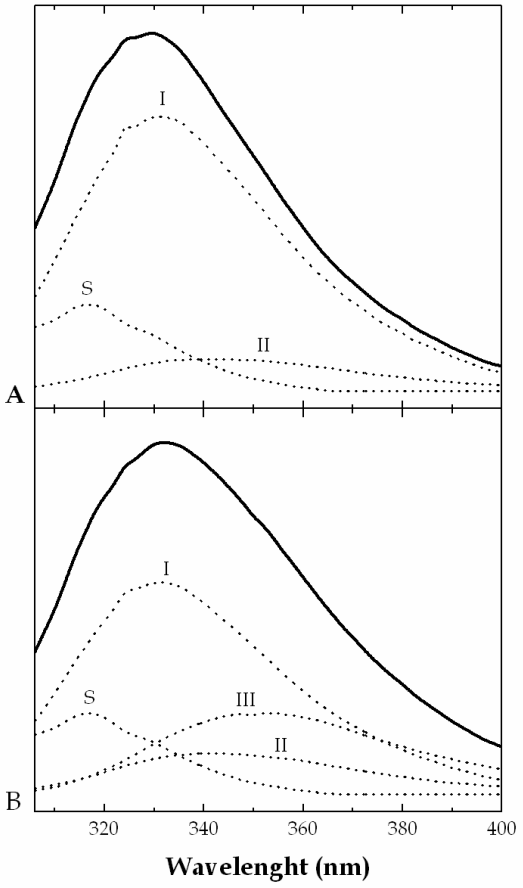

Figure 4.19: Analysis of the spectral components of the experimental fluorescence spectra of cardosin A in 10 % acetonitrile for 2 min (A) and for 170 h (B). Fitting was performed according to the theoretical model of discrete states of tryptophan residues in proteins (solid lines), which are the sums of the spectral components Tyr, S, I, II and III (dashed lines) (Section 3.6.6).

Figure 4.20A shows the CD spectra of cardosin A incubated in 10% acetonitrile, at pH 5, at different incubation times. The different curves clearly show that increasing incubation times modifies the spectrum: the CD spectrum of cardosin A at the beginning of the incubation is characteristic of all-β proteins (Venyaminov *et al*., 1996), and corresponds well to its native structure in water as shown previously (Oliveira, 2001). Upon incubation

of cardosin A with acetonitrile, the shape of the spectrum changed (Figure 4.20A), becoming very similar with a/β (a+β) proteins (Tanford, 1968) after 3 days of incubation.

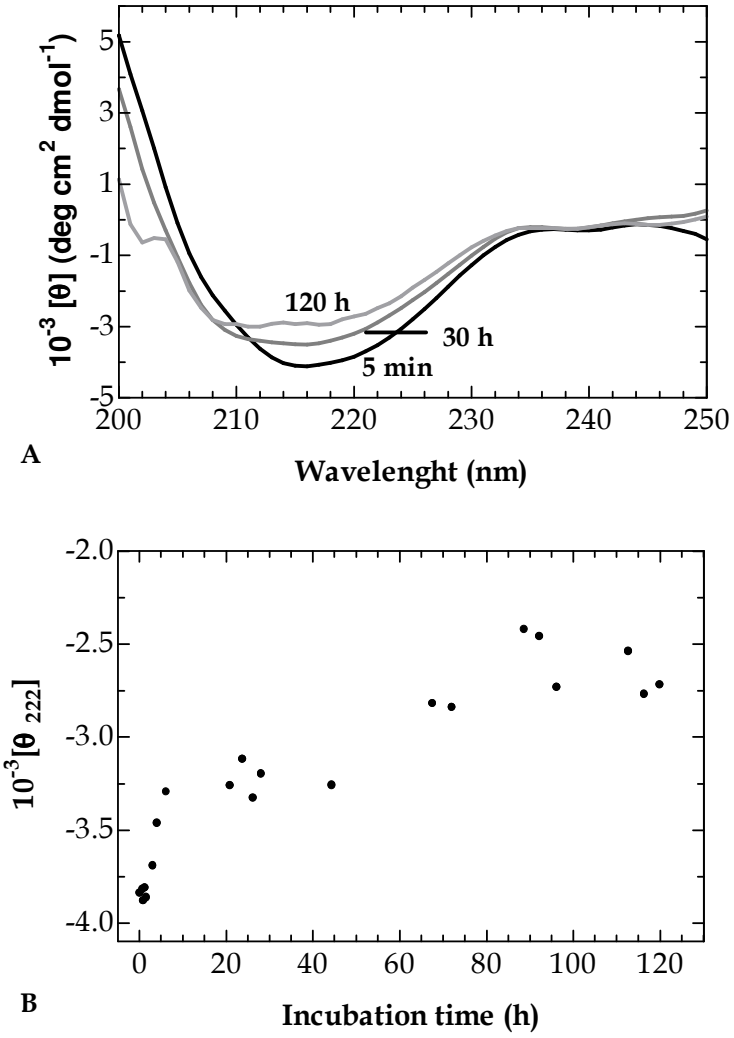

Figure 4.20: Far-ultraviolet CD monitoring of the conformational changes of cardosin A during incubation with 10 % acetonitrile. (A). Cardosin A CD spectra after some incubation times. (B) Ellipticity changes at 222 nm of cardosin A at various incubation times (Section 3.7.4).

To better emphasize the conformational changes induced by incubation with 10 % acetonitrile, transition curve has been constructed by recording the molar ellipticity at 222 nm (Figure 4.20B), where helical structures strongly absorb. As seen in Figure 4.20B, cardosin A in longer incubation times shows soft adjustments that indicate an increase in protein secondary structure. Overall signals for cardosin A in 10 % acetonitrile recorded at

222 nm show that longer acetonitrile incubation times lead to an increase in protein helicity, more pronounced up until 7 h, and after that less marked. As explained, these conformational changes, however, do not affect the heterodimeric association state and the molecular volume, as seen by SEC.

Some investigations have been reported that show time dependent induction of secondary structures of proteins in organic solvents, such as for acetonitrile (Tanford, 1968; Sato *et al*., 2000). In the studies performed with *a*-chimotrypsin the time-dependent alterations of secondary structure could explain the time dependent changes in catalytic activity. They suggested that native (distorted β-sheet) or native-like ($a+\beta$ type) structures are important for catalytic activity of this enzyme. Having this in mind, the formation of secondary structures in cardosin A was compared with the catalytic performance of the enzyme throughout the incubation process. Figure 4.21 shows the activity results obtained for cardosin A incubated at 25 ^{o}C in the presence of 10 % acetonitrile for up to 70 h. As seen in the previous chapter, for 1 h incubation in 10 % acetonitrile, there was an increase of catalytic activity. Likewise, in this experiment, after the addition of the solvent, activity increased by 30 % at least over the first 7 h as compared with the base level corresponding to the protein activity in pure buffer. This shows that the increase of cardosin A activity is directly related with the formation of secondary structure, ruling out the significance of substrate solvation in this catalytic performance. Additionally, after about 70 h of incubation with 10 % acetonitrile, activity returned to the initial base level.

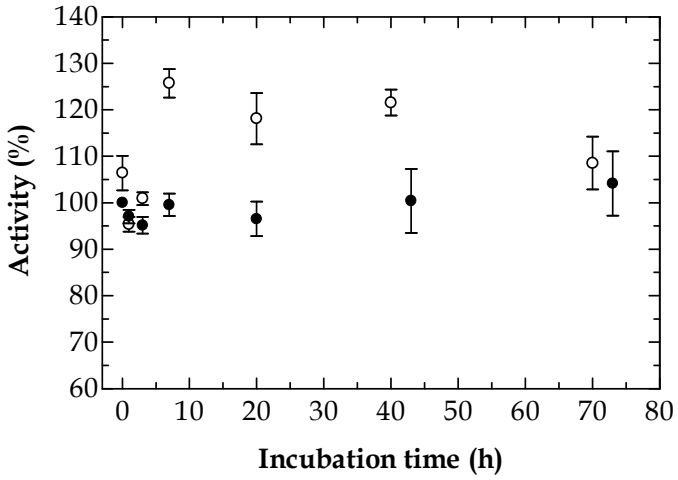

Figure 4.21: Effect of incubation time in cardosin A enzymatic activity. Cardosin A was incubated in buffer (closed circles) and in 10 % acetonitrile (v/v) (open circles) and activity determined according to Section 3.4.2.3.

Incubation with acetonitrile induced slow transformations of the protein that at a first stage increased activity and at a later stage stabilized it at the aqueous activity level. For longer incubation times, even though cardosin A displays native like activity its tertiary and secondary structures are not native. This indicates that this molecule can adopt different conformational states displaying higher activity or native like activity. Having in mind the thermodynamic analysis at pH 5 that showed different polypeptide chain stabilities in cardosin A it would be very interesting to know if these active acetonitrile induced conformational states in cardosin A are related with the different chain stabilities. In this sense thermodynamic analysis of the acetonitrile induced conformational states of cardosin A could shed light on this issue. Also the fact that these conformational states display native-like and enhanced activities constitutes an excellent opportunity for structure/functional analysis.

4.2.2.2 Thermodynamic analysis of acetonitrile induced conformational states

Due to the interest raised in these conformational states that display prominent activities, further characterization was envisaged. In this sense, the thermal stability of cardosin A was studied using high-sensitivity DSC at different incubation times with 10 % acetonitrile. Under all the experimental conditions employed, changing the scan rate from 90 to 12 ^{o}C/h afforded similar denaturation profiles, with transition temperatures differing less than 0.3 ^{o}C. Also, the reversibility of the cardosin A denaturation was nearly 85 % when the first temperature scan proceeded up to the temperatures at which the transition was 50% complete. Reversibility was reduced to 30 % when the transition reached 90 % completion. Thus, cardosin A denaturation could be approximated to a partial equilibrium and reversible process similarly to what was done for the thermodynamic analysis at several pH in aqueous buffer (Section 4.1.1.2). In consequence, the DSC data could be analyzed semi-quantitatively following thermodynamic models (Manly *et al*., 1985; Edge *et al*., 1985; Lin *et al*., 1994; Ruiz-Arribas *et al*., 1998; Kamen *et al*., 2000). The next step was to investigate the dependence of the thermal transition temperature of cardosin A denaturation versus the protein concentration in 10 % acetonitrile, as has been done previously in aqueous medium in Section 4.1.1.2. The thermal transition temperatures did not differ significantly (they differed by less than 0.3 ^{o}C) within the 8-90 µM concentration range. Taking into account that cardosin A is a heterodimeric protein, it may be concluded from this result that the thermal denaturation of cardosin A in the presence of acetonitrile is not correlated with the simultaneous dissociation of the folded dimer to the unfolded monomers similarly to what was verified for the thermal transition in aqueous conditions, as shown in Section 4.1.1.3.

It is clear that changes in the cardosin A structure during incubation of the protein in 10 % acetonitrile should alter all the characteristic thermodynamic functions of the protein denaturation transition. Figure 4.22, shows a set of the excess heat capacity functions for cardosin A obtained at different times of its incubation with 10 % acetonitrile. All traces were corrected for the instrumental baseline and the chemical baseline in accordance with (Takahashi *et al*., 1981). A first analysis show that the addition of 10 % acetonitrile to cardosin A solutions results in a decrease of the melting temperature by about 15 ^{o}C (from

about 70 °C for aqueous cardosin A to about 55 °C for cardosin A in 10 % acetonitrile). The results of deconvolution of the DSC profiles made under the assumption of the independent two-state model of unfolding are shown in the same figure and the estimated thermodynamic parameters presented in Table 4.7. It may be seen that the model correctly describes the thermal unfolding of cardosin A at different times of incubation with 10 % acetonitrile. The two-chain structure of cardosin A strongly indicates that the two-state model would correspond to the independent unfolding of the two chains of cardosin A heterodimer units, has it has been discussed before in Section 4.1.13.

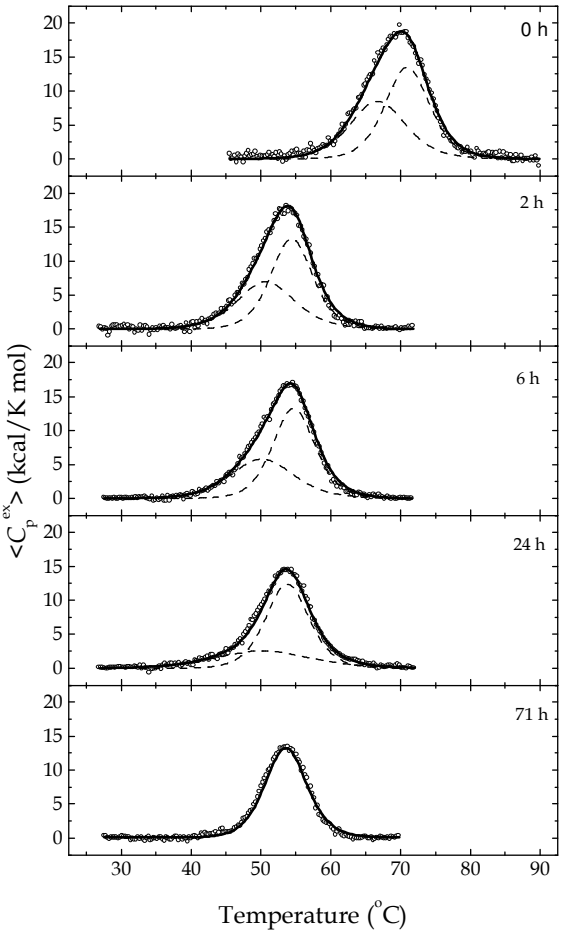

Figure 4.22: Temperature-dependence of the excess molar heat capacity of cardosin A (symbols) at the different times of incubation with 10 % acetonitrile. Experimental traces were corrected for the chemical baseline. Continuous lines are the result of non-linear least squares fittings of the experimental data of two independent two-state transitions (dashed lines) as implemented in the Origin software package (Section 3.8.2 and 3.8.3).

Table 4.7: Thermodynamic parameters for the individual transitions of cardosin A obtained by differential scanning calorimetry at different incubation times with 10 % acetonitrile at pH 5.0a.

Time of incubation (h)	First transition						Second transition					
	T_m (°C)	ΔH (T_m) (KJ/mol)	ΔC_p (KJ/°C mol)	ΔG^o (25°C) (KJ/mol)	T_s (°C)	ΔG^o (T_s) (KJ/mol)	T_m (°C)	ΔH (T_m) (KJ/mol)	ΔC_p (KJ/°C mol)	ΔG^o (25°C) (KJ/mol)	T_s (°C)	ΔG^o (T_s) (KJ/mol)
0	66.5	90.2	2.4	4.7	30.8	4.8	71.1	115.0	2.1	8.6	20.5	8.7
2	50.9	76.0	1.7	4.3	9.1	4.7	54.6	106.0	1.5	7.5	-9.0	10.7
6	50.4	69.5	-	-	-	-	54.7	106.0	-	-	-	-
24	50.4	47.0	-	-	-	-	53.9	104.0	-	-	-	-
71	-	-	-	-	-	-	53.6	105.8	-	-	-	-

[a] T_m is defined as the temperature at the midpoint of the unfolding transition (the standard deviation is ±0.2 °C); $\Delta H(T_m)$ is the calorimetric enthalpy of the unfolding transition with a standard deviation of ±5%; ΔC_p is the difference between the heat capacities of the intact and denatured states obtained from the slope of the graph of the temperature-dependence of $\Delta H(T_m)$ by pH variation of T_m (the standard deviations are ±0.2 kcal/°C mol); the free energy changes, ΔG^o, were calculated with the Gibbs-Helmholtz Equation 4; temperature of maximum stability, T_s, was calculated with equation 5.

Taking into account the marked overlapping of the thermal transitions for both chains, the only way to estimate the heat capacity changes, ΔCp, separately for both transitions is to exploit the linear relationship between the calorimetric enthalpies, ΔHcal, and the temperatures of the denaturation transitions, Tm. Experimentally, this can be achieved by varying the pH, as it has been shown for cardosin A solutions in pure buffer (Section 4.1.1.2). However, this method is not straightforwardly applicable to solutions containing acetonitrile owing to the kinetic nature of the structural changes in cardosin A in that solvent. Accordingly, it was possible to correctly estimate the ΔCp values using only the data for the first two hours of incubation of cardosin A in 10 % acetonitrile (Table 4.7). The calculated values of the temperatures of maximum stability, Ts, where the entropy change is zero, show that both chains of cardosin A change their hydrophobicity in the presence of 10 % acetonitrile: being relatively hydrophobic in water (Ts > 30 ^{o}C), they become relatively hydrophilic in water/acetonitrile mixtures (Ts < 10 ^{o}C) (Takahashi et al., 1981). This is a direct indication that acetonitrile destabilizes cardosin A by weakening the hydrophobic interactions between the non-polar residues of the protein. Examination of the free energy of protein stabilization calculated from the DSC data revealed that acetonitrile increases the thermodynamic stability of the large chain of cardosin A below room temperatures, in contrast to the destabilizing effect of acetonitrile at higher temperatures. The thermodynamic stability of the small chain of cardosin A decreased in all temperature regions, indicating the involvement of kinetic stabilization in the dynamics of the small chain. Comparing these results with the activity of cardosin A, it can be concluded that the unfolding of the small chain results in a decrease in enzymatic activity of cardosin A to the initial level, that is to say to the native-like activity.

4.2.2.3 General discussion

The time dependent induced changes in cardosin A in the presence of 10 % acetonitrile were investigated. A relatively slow kinetic effect of the presence of organic solvent on the conformation of the heterodimeric enzyme (cardosin A) allowed extensive monitoring of the physical-chemical parameters and structure changes, as well as measurements of enzymatic activity. All the spectroscopic data were consistent with soft changes in the secondary structure with no drastic changes in the higher-ordered structures of cardosin A. Similar results were obtained in previous works for enzymes incubated in water/organic solvent mixtures in a range from about 10 to 30 % volume ratio (Klibanov, 1997).

The thermal unfolding of cardosin A in 10 % acetonitrile at several incubations times were monitored by DSC and compared with thermal transition of cardosin A in pure buffer. All the thermograms showed that cardosin A in 10 % acetonitrile and at any incubation time is less stable. It is known that hydrophobic interactions play key roles in stabilizing the native conformations of proteins. A pronounced reduction in hydrophobic interactions due to almost any non-aqueous water miscible solvent must be intimately involved in the observed conformational changes. In the absence of such important interactions, the stability of a molecule is expected to be reduced, as seen with thermal transition of cardosin A in 10 % acetonitrile at any incubation time. Similar cases have been reported, for lysozyme in the presence of acetonitrile (Kovrigin *et al*., 2000), and for Ervatamin *C*, also in the presence of the same solvent (Sundd *et al*., 2004).

Additionally, calorimetric study of this system revealed an independent unfolding of the two chains of cardosin A unit. Thus, a stronger degree of destabilization was observed for one chain of cardosin A than for the other. This effect of acetonitrile can be interpreted in terms of an increase in the flexibility of the small chain of the enzyme. This can be substantiated by the enzymatic activity measurements results, where an increase in activity was observed during the first hours of incubation coinciding with the calorimetric data that suggests a small chain higher degree of destabilization (flexibility). However, a decrease in activity to the initial level was observed for the longest time of incubation, where total unfolding of one chain was observed. This indicates that conformational

rearrangements of the small chain do not preclude the catalytic function, and that soft destabilization can be directly related with activity enhancement. Therefore, solvent can indeed affect the enzymatic activity via conformational changes in the enzyme, as seen for the time dependent investigation of 10 % acetonitrile in cardosin A.

In Figure 4.19 both cardosin A in aqueous buffer and after 170 h incubation in 10 % acetonitrile fluorescence spectra were analyzed according to the model of discrete states of tryptophan residues in proteins. Analysis suggested that only one of the 5 tryptophan residues in cardosin A is moving to more solvent exposed environments. Having in mind the proposed small chain flexibility in such conditions and the cardosin A three dimensional structure, it can be hypothesized that Trp299 is changing its polar environment, since this tryptophan is the only present in the 15 kDa polypeptide chain, as seen in Figure 4.23. This residue is, as expected, not exposed to solvent and located in the middle of the 15 kDa chain. In this chain, it is positioned in the middle of a β–sheet symmetrically to Trp39, bordering the active site cleft (Frazão *et al*., 1999).

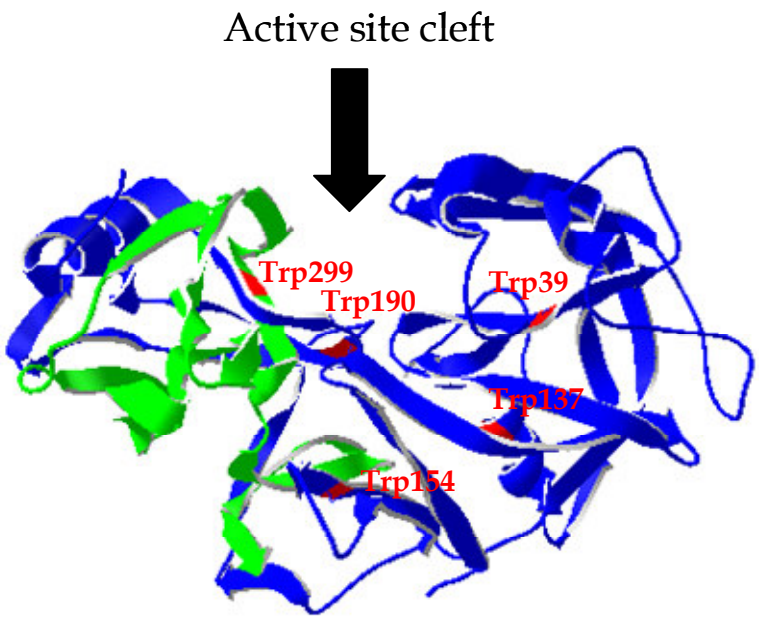

Figure 4.23: Cartoon representation of cardosin A three dimensional structure. Tryptophan residues in cardosin A are signaled in red. Residues in green represent the 15 kDa polypeptide chain and in blue the 31 kDa chain are highlighted. Cardosin A accession number is 1b5f.

The conformational flexibility of the small chain conformation has seen to occur in low acetonitrile concentration even enhancing its activity. Assuming that Trp299 movement inside the molecule is the main responsible for the fluorescence emission spectra seen before, and knowing that it is located just bordering the active site cleft, it is surprising to see the enhanced activity effects instead of deleterious ones. In order to understand the residues involved in substrate specificity for cardosin A, the mapping of substrate binding pockets was proposed a few years ago, as mentioned before (Frazão *et al.*, 1999). The structure of an inhibitor complexed with renin was fitted to the cardosin A coordinates and the docked structure energy minimized. Residues on each of the specificity sub-sites were defined as those having atoms within 4.0 Å of residues flanking the scissile peptide bond (Phe-Met). They included 127 atomic contacts including five putative hydrogen bonds involving several residues form cardosin A grouped according to their respective sub-sites and listed in Table 4.8. From this, two main observations can be drawn. First, that most of the cardosin A residues involved directly in catalysis belong to the 31 kDa

chain, and second, the only two residues from the 15 kDa chain, are Met289 and Ile300 (Table 4.8 in bold) , involved in close contacts with the substrate and close in sequence to Trp299. As can be seen in Figure 4.24A, residues Met289 and Ile300 apparently contact with the substrate and are in the primary structure vicinity of Trp299.

Table 4.8: List of cardosin A residues within 4.0 Å of the docked k-casein fragment and grouped with their sub-sites (Sn and S'n). Adapted from Frazão et al., 1999.

Subsite	Cardosin A residue	
S'3	Ser36	Ser37
	Ile73	Thr128
	Ile132	Tyr188
S'2	Gly34	Ser35
	Ile74	Tyr76
	Tyr189	
S'1	Gly76	Tyr181
	Phe213	Met289
	Ile300	
S1	Tyr75	Thr77
	Phe112	Phe117
	Ile120	Thr128
	Gly217	
S2	Gly76	Thr218
	Thr79	
S3	Ser13	Phe117
	Arg115	Ser219

Assuming that Trp299 is changing its position, becoming more exposed to the solvent and the enzyme remaining active, it can be hypothesised that the catalytic activity in cardosin A can still occur, and even be enhanced, when the small chain suffers conformational changes. This all point that these conformational changes most probably compromise Met289 and Ile300 residues close contacts with the active site.

Both Met289 and Ile300 were predicted to be in close contact with the substrate, more specifically as being part of the S'1 substrate binding pocket (Frazão *et al.*, 1999). The specific role and importance of each residue in substrate binding in cardosin A is still not completely known. Nevertheless, substrate specificity studies of cardosin A by peptide synthesis (Sarmento *et al.*, 2004a) suggested that S'1 binding pocket might be more hydrophobic and therefore with smaller volume, comparing with pepsin active site, due to replacement of residue 213, ILe in pepsin with Phe in cardosin A. Nevertheless, the specific role and involvement of the conserved Met289 and Ile300 is not defined. Up until now identification and characterization of residues important for catalysis has come from studying the specificity of pepsin, cathepsin D, plasmepsin II and others (Dunn, 2000). Overall, both S1 and S'1 subsites, neighbouring substrate binding pockects, are formed by large flat surfaces partially covered with loops producing shallow pockets. For S1 subsite the loop is the flexible flap from which the conserved Tyr75 and and Thr77 can interact with the substrate (Sielecki *et al.*, 1990; Okoniewska *et al.*, 1999). In light of the data available it is possible that Met289 and Ile300 interactions with substrate are not indispensable in hydrolysis or that in their absence (through conformational rearrangement of the small polypeptide chain) the interactions could be replaced by others. For instance, the important role of the flexible flap is well documented to be of major importance (Okoniewska *et al.*, 1999) and could prevail over changes at the S'1 subsite interaction, promoting the substrate / active site cleft interaction for peptide hydrolysis.

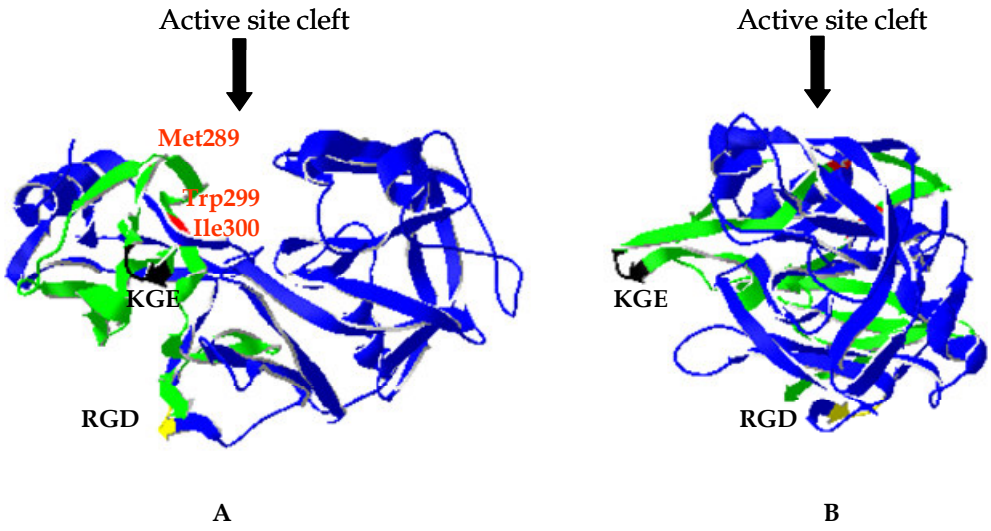

Figure 4.24: Cartoon representations of cardosin A three dimensional structure. A- Front view of cardosin A residues within 4 Å of the docked substrate (k-casein fragment) and belonging to the 15 kDa polypeptide chain are signaled in red. The single tryptophan residue in the 15 kDa chain is also signaled in red (Trp299). Residues in black and signaled as KGE represent the residues Lys278, Gly279 and Glu280, and in yellow represent the residues Arg176, Gly177 and Asp178. B- Left view of cardosin A with the KGE domain highlighted in black and RGD in yellow. Cardosin A accession number is 1b5f.

Recent work with cardosin A focused in the identification of phospholipase *Da* (PLD*a*) as the cardosin A-binding proteins and furthermore, it was described the involvement of the RGD (Arg176, Gly177 and Asp178) motif as well as the charge wise similar KGE sequence (Lys278, Gly279 and Glu280) in the interaction of these two plants proteins (Simões *et al*., 2005). In this investigation the complex formation determined between cardosin A and PLD*a* suggested possible concerted and/or synergistic actions in degenerative processes such as those observed during stress responses, plant senescence and/or pollen-pistil interactions. It was suggested that cardosin A association with PLD*a* may facilitate disintegration of the vacuoles in the dismantling phase of a vacuolar-type cell death. In this association the roles of RGD and KGE in complex formation were proposed. However how this could be accomplished in vivo still remains to be elucidated.

In Figure 4.24, RGD and KGE domains are signalled. It can be seen that RGD domain is located in the 31 kDa chain and KGE in the 15 kDa. The latter is located in a tip of a long

loop, like as RGD domain, in the other chain (Figure 4.24B). The KGE loop can be considered, therefore, a good candidate for belonging to a region of the small chain with high conformational flexibility. In fact, in the small chain, the KGE domain has the highest temperature *B* factors, 34.7, 76.0 and 44.7, respectively. When considering the results obtained for cardosin A in 10 % acetonitrile, it is reasonable to suggest that this flexible KGE region can be responsible for the secondary structures changes seen in 10 % acetonitrile cardosin A. Thus, in this situation, the concerted binding to PLD*a* with the RGD and KGE domains could be facilitated. Furthermore, our results show that in this complex formation, cardosin A could maintain its catalytic activity. Also, plant PLD*a* are involved in membrane degradation/lipid turnover during senescence or stress responses and also in signalling cascades (Wang, 2000). This all suggests that cardosin A and PLD*a* complex formation, and if both forms remained active, would allow for highly coordinated functions in vivo.

In conclusion, cardosin A molecule has some special features: its heterodimeric nature, where the catalytic aspartates are located in the same chain; on the other hand, each chain has its protein binding domains; and where some conformational freedom is allowed for the small chain without activity loss and that can even result in higher catalytic activity.

With this integrated approach, where acetonitrile induced effects were explored in detail it was possible to closely monitor not only the dramatic structural effects of acetonitrile, but also more subtle alterations. In these studies important information was revealed concerning the correlation between structure and function of heterodimeric cardosin A and extrapolations were made concerning in vivo physiological function.

5 Final Considerations

In this work project the folding of cardosin A with pH and with acetonitrile were studied and the induced conformational states characterized. A combination of techniques to monitor transitions associated with changes in pH and in acetonitrile was used. Intrinsic fluorescence and Far-UV CD, DSC, activity measurements and SEC experiments were carried out.

Cardosin A was showed to be active in the pH range 2.5-7.5, with maximum stability recorded at pH 5. In this range no conformational changes were detected and the activity seemed to be intimately associated with the protonated states of catalytic aspartates, as seen to occur with other APs. Unlike other APs that show narrower pH activity preferences that restrict them to a specific cellular compartment function, cardosin A could remain active in most cellular environments.

The thermal unfolding of cardosin A according to pH was seen to be partially reversible and non cooperative, characterised by independent unfolding of its polypeptide chains. The 15 kDa polypeptide chain was seen to unfold first, being less stable, whereas the 31 kDa chain was seen to be more stable.

The alkaline and denatured states were characterised and showed to be inactive and with different spectroscopic characteristics, as seen with other APs like pepsin, but neither corresponded to a fully unfolded protein. The acid induced conformational state was different from the native state and characterised by reduction of hydrodynamic volume of the molecule. On the other hand, the alkaline denatured state was seen to be more unfolded than the acid induced state and with wider molecular volume. This was seen to be associated with chains independent unfolding that, at higher pH resulted in chains dissociation.

Thermodynamic parameters determined so far for other AP are still far from sufficient in order to directly compare and assess APs stabilities. In this sense, the search for any relation between protein stability and the protein oligomeric state is still difficult to measure. On the other hand there is already some information that allows discussion of APs folding and function connection with the protein oligomeric state. In pepsin, monomeric AP, independent unfolding of the homologous domains was seen to occur, with N-terminal domain being less stable than the C-terminal and the physiological roles

of alkaline denatured states proposed in pepsin transportation to the stomach lumen. In this process, the control of pepsin activity occurs by different domain stabilities. In cardosin A, a heterodimeric counterpart, the different domains stabilities are replaced by different chain stabilities, with the small chain (corresponding to the C-terminal part of the protein) unfolding first. The heterodimeric nature of cardosin A and consequently its different thermal transition from pepsin has apparently no similar physiologic role, due to its activity and conformational stability in a wide pH range and where most physiological processes can take place. As for HIV-1 protease, the homodimeric counterpart, the unfolding is dictated by certain protein regions that display different stabilities and unfold sequentially. It was predicted that the flaps covering the active site cleft unfold first. In conclusion, the unfolding of APs studied so far seem to share the sequential unfolding of the molecule. More specifically, the sequential unfolding of domains, for pepsin, or of certain protein regions, for HIV-1 proteinase, or of polypeptide chains, for cardosin A. For pepsin, intimate involvements of unfolding events in physiological function were shown but for HIV-1 proteinase and cardosin A similar associations are still unknown.

Cardosin A physiological role has yet to be totally clarified as discussed before but it is possible that the heterodimeric nature of cardosin A, the different chains stabilities and the mentioned pH tolerance range are the result of an AP adaptation to multiple physiological functions in its specific environments. For example, the partial unfolding of the cardosin A can have physiological meaning, in the involvement in facilitating binding interactions with other protein domains, such as suggested for the interaction with phospholipase Dα C2 domain.

To investigate the existence of regulatory activity role of the small polypeptide chain of cardosin A, conformational states induced by acetonitrile in cardosin A were investigated. An integrated approach was done, where acetonitrile induced effects were explored in detail and the dramatic structural effects of acetonitrile and other more subtle alterations were described. Overall, acetonitrile showed two distinct effects in cardosin A. For higher acetonitrile concentrations (above 30 %) inactivation and drastic conformational changes were detected with no chain dissociation suggesting strong intersubunit interactions. In this situation acetonitrile apparently destabilizes cardosin A by non specific interactions in the molecule. On the other hand, low acetonitrile concentrations (1-10 %) enhanced

cardosin A activity accompanied by slight increase in secondary structure content. This effect was assumed to be the result of a direct effect of acetonitrile in the catalytic active site.

To further characterise cardosin A in low acetonitrile concentrations, time dependent induced changes in cardosin A in the presence of 10 % acetonitrile were investigated. A relatively slow kinetic effect of the presence of organic solvent on the conformation of the heterodimeric enzyme allowed extensive monitoring of the physical-chemical parameters and structure changes, as well as measurements of enzymatic activity. As expected the acetonitrile addition resulted in smaller stability as judged by thermodynamic parameters description. This system revealed also an independent unfolding of the two chains of cardosin A. A stronger degree of destabilization was observed for one chain of cardosin A than for the other. This effect of acetonitrile could be interpreted in terms of an increase in the flexibility of the small chain of the enzyme. Enzymatic activity measurements results, where an increase in activity was observed during the first hours of incubation coincided with the calorimetric data suggesting a small chain higher degree of destabilization (flexibility). Time dependent activity changes demonstrated that conformational rearrangements of the small chain did not preclude the catalytic function, and that soft destabilization could be directly related with activity enhancement.

In conclusion, cardosin A partially unfolded states displaying small chain conformational flexibility can occur without hampering activity. Furthermore, small chain flexibility could even enhance, therefore, regulate, cardosin A activity. In light of the acetonitrile induced effects in cardosin A and that partial unfolded states can be involved in facilitating binding processes with other protein domains, like in the interaction with phospholipase Dα C2 domain, coordinated functions in vivo can be, therefore, hypothesized.

6 Bibliography

Affleck, R, Xu, ZF, Suzawa, V, Focht, K, Clark, DS, Dordick, JS. 1992. Enzymatic catalysis and dynamics in low-water environments. Proc Natl Acad Sci USA. 89: 1100-1104.

Alexov, E. 2004. Numerical calculations of the pH of maximal protein stability. Eur J Biochem. 271: 173-185.

Andreeva, NS, Rumsh, LD. 2001. Analysis of crystal structures of aspartic proteinases: on the role of aminoacid residues adjacent to the catalytic site of pepsin-like enzymes. Prot Sci. 10: 2439-2450.

Anfinsen, CB, Scheraga, HA. 1975. Experimental and theoretical aspects of protein folding. Adv Prot Chem. 29: 205-301.

Barrett, A, Rawlings, ND, Woessner, JF. 1998. Handbook of proteolytic enzymes. Academic Press.

Batra, R, Gupta, MN. 1994. Enhancement of enzyme activity in aqueous-organic solvent mixtures. Biotechnol Lett. 16: 1059-1064.

Becktel, WJ, Schellman, JA. 1987. Protein stability curves. Biopolymers. 26: 1859-1877.

Beldarraín, A, Acosta, N, Montesinos, R, Mata, M, Cremata, J. 2000. Characterization of Mucor pusillus rennin expressed in Pichia pastoris: enzymic, spectroscopic and calorimetric studies. Biotechnol Appl Biochem. 31: 77-84.

Bell, G, Halling, PJ, Moore, BD, Partridge, J, Rees, DG. 1995. Biocatalyst behaviour in low-water systems. TIBTECH. 13: 468-473.

Bemquerer, MP, Adlercreutz, P, Tominaga, M. 1994. Pepsin-catalyzed peptide synthesis in organic media: studies with free and immobilized enzyme. Int J Pept Protein Res. 44(5): 448-56.

Berman, HM, Westbrook, J, Feng, Z, Gilliland, G, Bhat, TN, Weissig, H, Shindyalov, IN, Bourne, PE. 2000. The protein data bank. Nucleic Acid Research. 28: 235-242.

Bertie, JE, Lan, Z. 1997. Liquid water-acetonitrile mixtures at 25 C: The hydrogen-bonded structure studied through infrared absolute integrated absorption intensities. J. Phys. Chem. B. 101: 4111-4119.

Bielawski, JP, Yang, Z, 2000. Positive and negative selection in the DAZ gene family. Mol Biol Evol. 18, 523–528.

Blackburn, MN, Noltmann, EN. 1981. Evidence for an intermediate in the denaturation and assembly of phosphoglucose isomerase. Arch Biochem Biophys. 212: 162-169.

Bovey, F, Yanari, S. 1960. Pepsin. in Boyer, P, Lardy, H, Myrback, K, Eds. The enzymes, 2nd Ed. Academic Press, NY. Vol. 4, p 63.

Brandts, JF, Ju, CQ, Lin, L-N, Mas, MT. 1989. A simple model for proteins with interacting domains: Applications to scanning calorimetry data. Biochemistry. 28: 8588-8596.

Bromberg, LE, Klibanov, AM. 1995. Transport of proteins dissolved in organic solvents across biomimetic membranes. Proc Natl Acad Sci. 92: 1262–1266.

Brown, ED, Yada, RY. 1991. A kinetic and equilibrium study of the denaturation of aspartic proteinases from the fungi, Endothia parasitica and Mucor miehei. Biochim Biophys Acta 1076, 406-15.

Bruylants, G, Wouters, J, Michaux, C. 2005. Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design. Curr Med Chem. 12(17): 2011-20.

Burstein, EA. 1983. The intrinsic luminescence of proteins is a method for studies of the fast structural dynamics. Mol Biol (Moscow). 17: 455-467.

Burstein, EA. Vedenkina, NS, Ivkova, MN. 1973. Fluorescence and the location of tryptophan residues in protein molecules. Photochem Photobiol. 18: 263-279.

Butker, LG. 1979. Enzymes in non aqueous solvents. Enz Microb Technol 1: 253-259.

Campos, LA, Sancho, J. 2003. The active site of pepsin is formed in the intermediate conformation dominant at mildly acidic pH. FEBS Letters. 27038: 1-7.

Carginale, Trinchellab, F, Capassoa, C, Scudierob, R, Riggioa, M, Parisi, E. 2004. Adaptive evolution and functional divergence of pepsin gene family. Gene. 333: 81–90.

Carrea, G, Riva, S. 2000. Properties and Synthetic Applications of Enzymes in Organic Solvents. Angew Chem Int Ed. 39: 2226-2254.

Castanheira, P, Samyn, B, Sergeant, K, Clemente, JC, Dunn, BM, Pires, E, Beeumen, JV, Faro, C. 2005. Activation, proteolytic processing and peptide specificity of recombinant cardosin A. J Biol Chem. 280 (13): 13047-13054.

Castro, GR. 1999. Enzymatic activities of proteases dissolved in organic solvents Enzyme Microb. Technol. 25: 689–694.

Cataldo, AM, Nixon, RA. 1990. Enzymatically active lysosomal proteases are associated with amyloid deposits in Alzheimer brain. Proc Natl Acad Sci USA. 87: 3861-3865.

Chen, L, Erickson, JW, Rydel, TJ, Park, CH, Neidhart, D, Luly, J, Abad-Zapatero, C. 1992. Structure of a pepsin/renin inhibitor complex reveals a novel crystal packing induced by minor chemical alterations in the inhibitor Acta Crystallogr Sec B. 48: 476-488.

Chin, JT, Wheeler, SL, Klibanov, A. 1994. On protein solubility in organic solvents. Biotechnol Bioeng. 44: 140-145.

Cody, RJ. 1994. The clinical potential of rennin inhibitors and angiotensin antagonists. Drugs. 47: 586-598.

Cooper, A. 1998. Microcalorimetry of protein-protein interactions. in Ladbury, JE and Chowdhry, BZ, Eds. Biocalorimetry: the applications of calorimetry in the biological sciences. Wiley, p 103-111.

Cooper, JB. 2002. Aspartic proteinases in disease: a structural perspective. Current Drug Targets. 3: 155-173.

Cunningham, EL, Jaswal, SS, Sohl, JL, Agard, DA. 1999. Kinetic stability as a mechanism for protease longevity. Proc Natl Acad Sci. 96: 11008-11014.

D'Alessio, G. 1999. The evolutionary transition from monomeric to oligomeric proteins: tools, the environment, hypotheses. Prog Biophs Mol Biol. 72: 271-298.

D'Alessio. 1999b. Evolution of oligomeric proteins: The unusual case of a dimeric ribonuclease. Eur J Biochem. 266: 699-708.

D'Alessio. 2002. The evolution of monomeric and oligomeric βλ-type crystallins: facts and hypotheses. Eur J Biochem. 269: 3122-3130.

Dash, C, Kulkarni, A, Dunn, B, Rao, M. 2003. Aspartic peptidase inhibitors: implications in drug discovery.

Davies, DR. 1990. The structure and function of the aspartic proteinases. Annu Rev Biophys Chem. 19, 189-215.

Davis, MI, Douhéret, G. 1987. Comparisons of dielectric and thermodynamic ideality: II. Mixtures of associated liquids. Thermochim Acta 116: 183-194.

Dill, KA, Bromberg, S, Yue, K, Fiebig, KM, Yee, DP, Thomas, PD, Chan, HS. 1995. Principles of protein folding: A perspective from simple exact models. Prot Sci. 4: 561-602.

Domingos, A, Cardoso, PC, Xue Z-T, Clemente A, Brodelius, PE, Pais, MS. 2000. Purification, cloning and autoproteolytic processing of an aspartic proteinase from Centaurea calcitrapa. Eur J Biochem. 267: 6824-6831.

Drohse, HB, Foltman, B. 1989. Specificity of milk-clotting enzymes towards bovie k-casein. Biochem Biophys Acta. 995: 221-224.

Duarte, AS, Pereira, AO, Cabrita, AMS, Moir, AJG, Pires, EMV, Barros, MT. 2005. The characterisation of the collagenolytic activity of cardosin A demonstrates its potential application for extracellular matrix degradative processes. Curr Drug Discov Technol. 2: 37-44.

Duarte, AS. Cardosinas no estabelecimento de culturas primárias de neurónios. 2006. PhD thesis. University of Aveiro.

Duarte, AS, Rosa, N, Duarte, EP, Pires, E, Barros, MT. 2006. Cardosins: a new and efficient plant enzymatic tool to dissociate neuronal cells for the establishment of cell cultures. Biotechnol Bioeng. Accepted for publication.

Dunn, BM, Hung, S-H. 2000. The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases. Biochim Biophys Acta. 1477: 231-240.

Dunn, BM. 2002. Structure and mechanism of the pepsin-like family of aspartic peptidases. Chem Rev. 102: 4431-4458.

Durham, TB, Shepherd, TA. 2006. Progress toward the discovery and development of efficacious BACE inhibitors. Curr Opin Drug Discov Devel. 9(6): 776-91.

Eder, J, Fersht, AR. 1995. Pro-sequence-assisted protein folding. Mol Microbiol. 16: 609-614.

Edge, V, Allewell, NM, Sturtevant, JM. 1985. High resolution differential scanning calorimetric analysis of the subunits of Escherichia coli aspartate transcarbamoylase. Biochemistry. 24: 5899-5906.

Emi, S, Myers, DV, Iacobucci, GA. 1976. Purification and properties of the thermostable acid protease of Penicillium duponti. Biochemistry. 15: 842-848.

Fadnavis, NW, Seshadri, R, Sheelu, G, Madhuri, KV. 2005. Relevance of Frank's solvent classification as typically aqueous and typically non-aqueous to activities of firefly luciferase, alcohol dehydrogenase, and alpha-chymotrypsin in aqueous binaries. Arch Biochem Biophys 433: 454-465.

Faro, CJ, Moir, AGJ, Pires, EMV. 1992. Specificity of a milk clotting enzyme extracted from the thistle Cynara cardunculus L: action on oxidised insulin and k-casein. Biotechnol Lett 14(9): 841-846.

Faro, C, Ramalho-Santos, M, Vieira, M, Mendes, A, Simões, I, Andrade, R, Veríssimo, P, Lin, X, Tang, J, Pires, E. 1990. Cloning and characterization of cDNA encoding cardosin A, an RGD-containing plant aspartic proteinase. J Biol Chem. 274: 28724-28729.

Faro, C, Veríssimo, P, Lin, Y, Tang, J, Pires, E. 1995. Cardosin A and B, aspartic proteases from the flowers of cardoon. Adv Exp Med Biol. 362: 373-377.

Favilla, R, Parisoli, A, Mazzini, A. 1997. Alkaline denaturation and partial refolding of pepsin investigated with DAPI as an extrinsic probe. Biophys Chem. 67, 75-83.

Fersht, A. 1999. Basic equations of Enzyme Kinetics. In Freeman, WH, Ed. Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding. USA. Chp 3, pp 103-131.

Fitzpatrick, PA, Steinmetz, ACU, Ringe, D, Klibanov, AM. 1993. Enzyme crystal structure in a neat organic solvent. Proc Nat Acad Sci USA 90: 8653-8657.

Frazão, C, Bento, I, Costa, J, Soares, CM, Veríssimo, P, Faro, C, Pires, E, Cooper, J, Carrondo, MA. 1999. Crystal structure of cardosin A, a glycosilated and Arg-Gly-Asp-contaning aspartic proteinase from the flowers of Cynara cardunculus L. J. Biol. Chem. 274 (39): 27694-27701.

Fruton, JS. 1971. Pepsin. in Boyer, PD, Ed. The Enzymes, Academic Press, New York Vol. III, p. 119-164.

Fruton, JS. 2002. A History of pepsin and related enzymes. The Quarterly Review of Biology. 77: 127–147.

Fukada, H, Sturtevant, JM, Quiocho, FA. 1983. Thermodynamics of the Binding of L-Arabinose and of D-Galactose to the L-Arabinose-binding protein of Escherichia coli. J Biol Chem 258: 13193-13198.

Gagnon-Arsenault, I, Tremblay, J, Bourbonnais, Y. Fungal yapsins and cell wall: a unique family of aspartic peptidases for a distinctive cellular function. FEMS Yeast Res. 6: 966-978.

Grant, SK, Deckman, IC, Culp, JS, Minnich, MD, Brooks, IS, Hensley, P, Debouck, C, Meek, TD. 1992. Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases. Biochem. 31, 9491-9501.

Greenfield, N, Pasman, GD. 1969. Computed circular dichroism spectra for the evaluation of protein conformation. Biochem. 8: 4108-4116.

Griebenow, K, Klibanov, AM. 1996. On protein denaturation in aqueous-organic mixtures but not in pure organic solvents. J Am Chem Soc. 118 (47): 11695-11700.

Griko, YV, Venyaminov, SY, Privalov, PL. 1989. Heat and cold denaturation of phosphoglycerate kinase (interaction of domains). FEBS Lett. 244: 276-278.

Guagliardi, A, Manco, G, Rossi, M, Bartolucci, S. 1989. Stability and actvity of a thermostable malic enzyme in denaturants and water-miscible solvents. Eur J Biochem. 183:25-30.

Guillaume, YC, Guinchard, C. 1997. ACN clusters in a water/ACN mixture, with implications for the RPLC weak polar solute retention. Anal. Chem. 69: 183-189.

Gupta, MN, Batra, R, Tyagi, R, Sharma, A. 1997. Polarity Index: The Guiding Solvent Parameter for Enzyme Stability in Aqueous-Organic Cosolvent MixturesBiotechnol. Progr. 13: 284–288.

Hong, L, Koelsch, G, Lin, XL, Wu, SL, Terzyan, S, Ghosh, AK, Zhang, XC, Tang, J. 2000. Structure of the protease domain of memapsin-2 complexed with inhibitor. Science. 290: 150-153.

Hori, H, Yoshino, T, Ishisuka, Y, Yamauchi, T, Murakami, K. 1988. Role of N-linked oligosaccharides attached to human renin expressed in COS cells. FEBS Lett. 232: 391-394.

Hyland, LJ, Tomaszek, TA, Meek, TD. 1991. Human immunodeficiency virus-1 protease: 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. Biochemistry. 30: 8454-8463.

Ido, E, Han, HP, Kezdy, FJ, Tang, J. 1991. Kinetic studies of human immunodeficiency virus type 1 protease and its active-site hydrogen bond mutant A28S. J Biol Chem. 266, 24359-24366.

Jaenicke, R. 1987. Folding and association of proteins. Prog Byophys Mol Biol. 49: 117-237.

James, MNG, Sielecki, RG. 1985. Stereochemical analysis of peptide bond hydrolysis catalyzed by the aspartic proteinase penicillopepsin. Biochemistry. 24: 3701-3713.

Jones, S, Thornton, JM. 1995. Protein-protein interactions: a review of protein dimer structures. Prog Biophys Mol Biol 63: 31-65.

Jones, S, Marin, A, Thornton, JM. 2000. Protein domain interfaces: characterization and comparison with oligomeric protein interfaces. Prot Eng. 13 (2): 77-82.

Kamatari, YO, Dobson, CM, Konno, T. 2003. Structural dissection of alkaline-denatured pepsin. Prot Sci. 12: 717-724.

Kamen, DE, Griko, Y, Woody, RW. 2000. The stability, structural organization, and denaturation of pectate lyase C, a parallel beta-helix protein. Biochemistry 39: 15932-15943.

Kempner, ES. 1993. Movable lobes and flexible loops in proteins. Structural deformations that control biochemical activity. FEBS Letts. 326: 4-10.

Kenji, T, Senarath, A, Koji, M, Sanath, R, Masayuki, K, Masaki, K, Nobuko, K-H, Akihiro, I, Chiaki, S, Hideshi, I. 2005. Nepenthesin, a unique member of a novel sub-family of aspartic proteinases: enzymatic and structural characteristics. Curr Prot Pep Sci. 6(6): 513-525.

Kervinen, J, Tobin, GJ, Costa, J, Waugh, DS, Wlodaver, A, Zdanov, A. 1999. Crystal structure of plant aspartic proteinases prophytepsin:inactivation and vacuolar targeting. EMBO J. 18 (14): 3947-3955.

Khan, AR, James, MNG. 1998. Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. Prot Sci. 7: 815-836.

Khan, RH, Rasheedi, S, Haq, SK. 2003. Effect of pH, temperature and alcohols on the stability of glycosylated and deglycosylated stem bromelain. J. Biosci. 28 (6). 709–714.

Kijima, T, Yamamoto, S, Kise, H. 1996. Study on tryptophan fluorescence and catalytic activity of α-chymotrypsin in aqueous-organic media. Enz Microb Technol. 18: 2-6.

Klibanov, A.M. 1997. Why are enzymes less active in organic solvents than in water? Trends Biotechnol. 15: 97-101.

Knight, CG. 1995. Active-site titration of peptidases. in Meth Enzymol. Academic Press. Vol. 248, Chapter 6, 85-101.

Koelsch, G, Mares, M, Metcalf, P, Fusek, M. 1994. Multiple functions of pro-parts of aspartic proteinases zymogens. FEBS Lett. 343: 6-10.

Kovrigin, EL, Potekhin, SA. 2000. On the stabilizing action of protein denaturation: acetonitrile effect on stability on stability of lysozyme in aqueous solution. Biophys Chem. 83: 45-59.

Krausslich, HG, Winner, E. 1988. Viral Proteinases. Annu Rev Biochem. 57: 701-754.

Laemmli, UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227:-680.

Lah, T, Drobnic-Kosorok, M, Turk, V, Pain, RH. 1984.Conformation, structure and activation of bovine cathepsin D. Unfolding and refolding studies. Biochem J. 218: 601-608.

Ledgerwood, EC, Brennan, SO, Cawley, NX, Loh, YP, George, PM. 1996. Yeast aspartic protease 3 (Yap3) prefers substrates with basic residues in the P2, P1 and P2' positions. FEBS Lett. 383(1-2): 67-71.

Lijn, RV, Janssen, AEM, Moore, BD, Halling, PJ, Nelly, SM, Price, NC. 2002. Reversible acetonitrile-induced inactivation/activation of thermolysin. Chem Bio Chem. 3: 1112-1116.

Lin, L-N, Mason, AB, Woodworth, RC, Brandts, JF. 1994. Calorimetric studies of serum transferrin and ovotransferrin. Estimates of domain interactions, and study of the kinetic complexities of ferric ion binding. Biochemistry 33: 1881-1888.

Lin, X, Loy, JA, Sussman, F, Tang, J. 1993. Conformational instability of the N- and C-terminal lobes of porcine pepsin in neutral and alkaline solutions. Prot Sci. 2: 1383-1390.

Lin, Y, Fusek, M, Lin, X, Hartsuck, JA, Kezdy, FJ, Tang, J. 1992. pH dependence of kinetic parameters of pepsin, rhizopuspepsin, and their active-site hydrogen bond mutants. J Biol Chem. 267 (26): 18413-18418.

Lopes, HJA. 2003. Caracterização de uma proteinase aspártica de Cynara cardunculus L. por electroforese em gel e espectrometria de massa. Msc Thesis. University of Aveiro. Portugal.

Luo, P, Baldwin, RL. 1997. Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. Biochemistry. 36: 8413–8421.

Makowski, GS, Ramsby, ML. 1997. In: Creighton T.E., Eds. Protein structure. A practical approach. Oxford University Press, Oxford, p 1-28.

Manly, SP, Mathews, KS, Sturtevant, JM. 1985. Thermal denaturation of core protein of lac repressor. Biochemistry. 24: 3842-3846.

Marcus, Y, Migron, Y. 1991. Polarity, hydrogen bonding, and structure of mixtures of water and cyanomethane. J Phys Chem. 95: 400-406.

Mathews, RR. 1993. Structural and Genetic Analysis of Protein Stability. Annu. Rev. Biochem. 62: 653-683.

McPhie, P. 1989. A reversible unfolding reaction of swine pepsin; implications for pepsinogen's folding mechanism Biochem Biophys Res Commun. 158, 115-119.

Munson, M, Balasubramanian, S, Fleming, KG, Nagi, AD, O'Brian, R, Sturtevant, JM, Regan, L. 1996. What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. Prot Sci. 5: 1584-1593.

Mutlu, A, Pfeil, JE, Gal, S. 1999. A probarley lectin processing enzyme purified from Arabidopsis thaliana seeds. Phytochemistry. 47(8): 1453-1459.

Nakanishi, K, Matsuno, R. 1986. Kinetics of enzymatic synthesis of peptides in aqueous/organic biphasic systems. Thermolysin-catalyzed synthesis of N-(benzyloxycarbonyl)-L-phenylalanyl-L-phenylalanine methyl ester. Eur J Biochem. 161: 533-540.

Neet, KE, Lee, JC. 2002. Biophysical characterization of proteins in the post-genomic era of proteomics. Mol Cell Proteomics. 1: 415-420.

Neet, KE, Timm, DE. 1994. Conformational stability of dimeric proteins: quantitative studies by equilibrium denaturation. Prot Sci. 3: 2167-2174.

Northrop, JH. 1930. Crystalline pepsin. I. Isolation and tests of purity. J Gen Physiol. 13: 739-766.

Ogino, H, Watanabe, F, Yamada, M, Nakagawa, S, Hirose, T, Noguchi, A, Yasuda, M, Ishikawa, H. 1999. Purification and characterization of organic solvent-stable protease from organic solvent-tolerant Pseudomonas aeruginosa PST-01. J Biosci Bioeng. 87: 61–68.

Ohtsuru, M, Tang, J, Delaney, R. 1982. Purification and characterization of rhizopuspesin isozymes from a liquid culture of Rhizopus chinensis. Int J Biochem. 14: 925-932.

Okoniewska, M, Tanaka, T, Yada, RY. 1999. The role of the flap residue, threonine 77, in the activation and catalytic activity of pepsin A. Prot Eng. 12 (1): 55-61.

Oliveira, C. 2001. Efeito dos solventes orgânicos na estrutura da cardosina A. Msc Thesis. University of Aveiro, Portugal.

Oyama, H, Abe, S, Ushiyama, S, Takahashi, S, Oda, K. 1999. Identification of catalytic residues of pepstatin-insensitive carboxyl proteinases from prokaryotes by site-directed mutagenesis. J Biol Chem. 274(39): 27815-27822.

Pace, CN. 1990. Conformational stability of globular proteins. Trends Biochem. Sci. 15: 14-17.

Pace, CN, Shirley, BA, Mcnutt, M, Gajiwala, K. 1996. Forces contributing to the conformational stability of proteins. FASEB. 10: 75-83.

Parida, S, Dordick, JS. 1991. Substrate Structure and Solvent Hydrophobicity Control Lipase Catalysis in Organic Media. J Am Chem Soc. 113: 2253-2259.

Pearl, LH, and Blundell, TL. 1984. The active sites of aspartic proteinases. FEBS Lett. 174: 96-101.

Pereira, AO, Cartucho, DJ, Duarte, AS, Gil, MH, Cabrita, AMS. 2005. Immobilisation of Cardosin A in chitosan sponges as a novel implant for drug delivery. Curr Drug Discov Technol. 2: 231-238.

Pfeil, W. 1998. Protein stability and folding: a collection of thermodynamic data. in Rheinau, CM, Ed. Springer-Verlag. Berlin, Heidelberg.

Pichova, I, Pavlikova, L, Dostal, J, Dolej, E, Hru, K, Heidingsfeldova, O, Weber, J, Ruml, T, Souek, M. 2001. Secreted aspartic proteases of C. albicans, C. tropicalis, C. parapsilosis and C. lusitaniae. Inhibition with peptidomimetic inhibitor. Science. 290: 150-153.

Privalov, PL. 1990. Cold denaturation of proteins. Crit Rev Biochem. Mol Biol. 25(4): 281–305.

Privalov, PL. 1979. Stability of proteins: Small globular proteins. Adv Prot Chem. 33: 167-239.

Privalov, PL. 1982. Stability of proteins: Proteins which do not present a single cooperative system. Adv Prot Chem. 35: 1-104.

Privalov, PL, Khechinashvili, NN. 1974. A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. J Mol Biol. 86: 665-684.

Privalov, PL, Mateo, PL, Khechinashvili, NN, Stepanov, VM, Revina, LP. 1981. Comparative thermodynamic study of pepsinogen and pepsin structure. J Mol Biol. 152, 445-464.

Radlowski, M. 2005. Proteolytic enzymes from generative organs of flowering plants (Angiospermae). J Appl Genet. 46 (3): 247-257.

Rajagopalan, TG, Stein, WH, Moore, S. 1966. The inactivation of pepsin by diazoacetyl-DL-norleucine methyl ester. J Biol Chem. 241: 4295-4297.

Ramalho-Santos, M, Pissarra, J, Veríssimo, P, Pereira, S, Salema, R, Pires, E, Faro, CJ. 1997. Cardosin A, an abundant aspartic proteinase, accumulates in protein storage vacuoles in the stigmatic papillae of Cynara cardunculus L. Planta. 203: 204-212.

Ramalho-Santos, M, Veríssimo, P, Cortes, L, Samyn, B, van Beeumen, J, Pires, E. 1998. Identification and proteolytic processing of procardosin A. Eur J Biochem. 255: 133-138.

Ramos CH, Ferreira ST. 2005. Protein folding, misfolding and aggregation: evolving concepts and conformational diseases. Protein Pept Lett. 12(3): 213-22.

Rao, JK, Erickson, JW, Wlodawer, A. 1991. Structural and evolutionary relationships between retroviral and eucaryotic aspartic proteinases. Biochemistry. 30(19): 4663-71.

Raymond, MN, Bricas, E, Salesse, R, Garnier, J, Garnot, P, Ribadeau Dumas, B. 1973. A proteolytic unit for chymosin (rennin) activity based on a reference synthetic peptide. J Dairy Sci. 56: 419-422.

Rawlings, ND, Barrett, AJ. 1995. Families of aspartic peptidases, and those of unknown catalytic mechanism. Methods Enzymol. 248: 105-120.

Rawlings, ND, Morton, FR, Barrett, AJ. 2006. MEROPS: the peptidase database. Nucleic Acids Res 34: 270-272.

Reshetnyak, YK, Burstein, EA. 2001. Decomposition of protein tryptophan fluorescence spectra into log-normal components. II. The statistical proof of discreteness of tryptophan classes in proteins. Biophys J. 81 (3): 1710- 1734.

Ritcher, C, Tanaka, T, Yada, RY. 1998. Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. Biochem J. 335: 481-490.

Rochefort, H. 1990. Biological and clinical significance of cathepsin D in breast cancer. Semin Cancer Biol. 1: 153-160.

Ruiz-Arribas, A, Zhadan, GG, Kutyshenko, VP, Santamaría, RI, Cortijo, M, Villar, E, Fernández-Abalos, JM, Calvette, JJ, Shnyrov, VL. 1998. Thermodynamic stability of two variants of xylanase (Xys1) from Streptomyces halstedii JM8. Eur J Biochem. 253 (1998) 462-468.

Rupley, JA, Careri, G. 1991. Protein hydration and function. Adv Prot Chem 41, 37-172.

Sakurai, T, Margolin, AL, Russel, AJ, Klibanov, AM. 1988. Control of enzyme enantioselectivity by the reaction medium. J Am Chem Soc. 110: 7236-7237.

Sarkkinen, P, Kalkkine, N, Tilgmann, C, Siuro, J, Kervinen, J, Mikola, L. 1992. Aspartic proteinase from barley grains is related to mammalian lysossomal cathepsin D. Planta. 186: 317-323.

Sarmento, AC. 2002. A cardosina A uma proteinase aspártica vegetal como modelo para estudos de comportamento em solventes orgânicos. Da biocatálise à química de proteínas. PhD thesis. University of Aveiro.

Sarmento, AC, Oliveira, CS, Duarte, AS, Pires, E, Barros, M. 2006. Evaluation of cardosin A as a probe for limited proteolysis in non-aqueous environments—complex substrates hydrolysis. Enz Microb Technol. 38: 415–421.

Sarmento, AC, Oliveira, CS, Pereira, A, Esteves, VI, Moir, A, Halling, PI, Saraiva, J, Pires, E, Barros, M. 2006. Characterization of cardosin A unfolding in organic solvents. Submited to FEBS J.

Sarmento, AC, Oliveira, CS, Pires, E, Amado, F, Barros, M. 2004a. Reverse hydrolysis by cardosin A: specificity considerations. J Mol Cat B: Enzymatic. 28: 33-37.

Sarmento, AC, Oliveira, CS, Pires, E, Halling, P, Barros, M. 2003. Cardosin A as a model aspartic proteinase for the study of organic solvent effects. An overview on catalytic and structural aspects. J Mol Cat B: Enzymatic. 21: 19-23.

Sarmento, AC, Oliveira, CS, Pires, E, Halling, P, Barros, M. 2004b. Evaluation of cardosin A as a proteolytic probe in the presence of organic solvents. J Mol Cat B: Enzymatic. 31: 137-141.

Sarmento, AC, Silvestre, L, Barros, M, Pires, E. 1998. Cardosins A and B, two new enzymes available for peptide synthesis. J Mol Cat B: Enzymatic. 5: 327-330.

Sato, M, Sasaki, T, Kobayashi, M, Kise, H. 2000. Time-dependent structure and activity changes of α-chymotrypsin in water/alcohol mixed solvents. Biosci Biotech Biochem 64: 2552-2558.

Schmid, F. 2005. Spectroscopic techniques to study protein folding and stability. In: Buchner, J, Kiefhaber, T. eds. Protein folding handbook. Part I. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. Chapter 2.

Seelmeier, S, Schmidt, H, Turk, V, Der Helm, K. 1988. Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. Proc Natl Acad Sci U S A. 85: 6612-6616.

Schmid, FX. 1997. Optical spectroscopy to characterize protein conformation and conformational changes. in Creighton TE. Ed. Protein structure: a practical approach. IRL Press, Oxford. Chp 1.

Sielecki, AR, Fedorov, AA, Boodhoo, A, Andreeva, NS, James, MNG. 1990. Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 Å resolution. J Mol Biol. 214: 143-170.

Silva, SV, Malcata, FX. 2005. Partial Identification of Water-Soluble Peptides Released at Early Stages of Proteolysis in Sterilized Ovine Cheese-Like Systems: Influence of Type of Coagulant and Starter. J Dairy Sci. 88:1947–1954.

Silva, AM, Lee, AY, Gulnik, SV, Majer, P, Collins, J, Bhat, TN, Collins, PJ, Cachau, RE, Luker, KE, Gluzman, IY, Francis, SE, Oksman, A, Goldberg, DE, Erickson, JE. 1996. Structure and inhibition of plasmepsin II, a haemoglobin degrading enzyme from Plasmodium falciparum. Proc Natl Acad Sci USA. 93: 10034-10039.

Simões, I, Faro, C. 2004. Structure and function of plant aspartic proteinases. Eur J Biochem. 271: 2067-2075.

Simões, I, Mueller, E, Otto, A, Bur, D, Cheung, AY, Faro, C, Pires, E. 2005. Molecular analysis of the interaction between cardosin A and phospholipase Dα. FEBS Journal. 272: 5786-5798.

Singh, SM, Panda, AK. 2005. Solubilization and refolding of bacterial inclusion body of proteins. J Biosci Bioeng. 99(4): 303-10.

Seoighe, C, Johnston, CR, Shields, DC, 2003. Significantly different patterns of amino acid replacement after gene duplication as compared to after speciation. Mol Biol Evol. 20: 484– 490.

Suguna, K, Padlan, EA, Smith, KW, Carlson, WD, Davies, DR. 1987. Binding of a reduced peptide inhibitor to the aspartic proteinases from Rhizopus chinensis: implications for a mechanism of action. Proc Natl Acad Sci USA. 84(20): 7009-13.

Sundd, M, Kundu, S, Dubey, VK, Jagannadham, MV. 2004. Unfolding of ervatamin C in the presence of organic solvents: sequential transitions of the protein in the O-state. J Biochem Mol Biol. 37(5): 586-96.

Szeltner, Z, Polgar, L. 1996. Conformational stability and catalytic activity of HIV-1 protease are both enhanced at high salt concentration J Biol Chem. 271, 5458- 5463.

Takahashi, K, Sturtevant, JM. 1981. Thermal denaturation of Streptomyces subtilisin inhibitor, subtilisin BPN', and the inhibitor-subtilisin complex. Biochemistry. 20: 6185-6190.

Tanford, C. 1968. Protein denaturation. Adv Protein Chem. 23: 121-282.

Tang, J. 1971. Specific and irreversible inactivation of pepsin by substrate-like epoxides. J. Biol. Chem. 246: 4510-4517.

Tang, J, Wong, NS. 1987. Evolution in the structure and function of aspartic proteases. J Cell Biochem. 33: 53-63.

Tcherkasskaya, O, Ptitsyn, OB. 1999. Molten globule versus variety of intermediates: influence of anions on pH-denatured apomyoglobin. FEBS Letters. 455(3): 325-331.

Todd, MJ, Semo, N, Freire, E. 1998. The structural stability of the HIV-1 protease. J Mol Biol. 283: 475-488.

Uversky, VN. 1997. Diversity of equilibrium compact forms of denatured globular proteins. Prot Pept Lett. 4: 355-367.

Uversky, VN, Segels, DJ, Doniach, S, Fink, AL. 1998. Association-induced folding of globular proteins. Proc Natl Acad Sci USA. 95: 5480–5483.

Uversky, VN. 2002. Cracking the folding code. Why do some proteins adopt partially folded conformations, whereas others don't? FEBS Letters. 514: 181-183.

van Erp, SHM, Kamenskaya, EO, Khmelmtsky, YL. 1991. The effect of water content and nature of organic solvent on enzyme activity in low-water media. A quantitative description. Eur J Biochem. 202: 379-384.

Vazquez-Duhalt, R, Semple, KM, Westlake, DWS, Fedorak, PM. 1983. Effect of water-miscible organic solvents on the catalytic activity of cytochrome c. Enz Microb Technol. 15: 936-943.

Veerapandian, B, Cooper, JB, Sali, A, Blundell, TL, Rosati, RL, Dominy, BW, Damon, DB, Hoover, DJ. 1992. Direct observation by X-ray analysis of the tetrahedral "intermediate" of aspartic proteinases. Prot Sci. 1, 322-328.

Venyaminov, SY, Yang, JT. 1996. Circular dichroism and the conformational analysis of biomolecules. in Fasman GD, Ed. Plenum Press, New York p 69-107.

Veríssimo, P, Faro, C, Moir, JG, Ling, Y, Tang, J, Pires, E. 1996. Purification, characterization and partial amino acid sequencing of two new aspartic proteinases from fresh flowers of Cynara cardunculus L. Eur J Biochem. 235: 762-768.

Volkin, DB, Klibanov, AM. 1991. Alterations in the structure of the proteins that cause their irreversible inactivation. Develop Biol Standard. 74: 73-81.

Wang, X. 2000. Multiple forms of phospholipase D in plants: the gene family, catalytic and regulatory properties, and cellular functions. Prog Lipid Res. 39: 109-149.

Whitaker, JR 1963. Determination of molecular weights of proteins by gel filtration on Sephadex. Anal Chem. 35: 1950-1953.

Whitaker,JR. 1970. Protease of Endothia parasitica. Meth Enzymol. 19: 436-445.

Williams, MA, Goodfellow, JM, Thornton, JM. 1994. Buried waters and internal cavities in monomeric proteins. Prot Sci. 3: 2167-2174.

Wittlin, S, Rosel, J, Hofmann, F, Stover, DR. 1999. Mechanisms and kinetics of procathepsin D activation. Eur J. Biochem. 265: 384-393.

Wlodawer, A, Vondrasek, J. 1998. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. Ann Rev Biophys Bioeng. 27: 249-284.

Wray, W, Boulikas, T, Wray, VP, Hancock, R. 1981. Silver staining of proteins in polyacrylamide gels. Anal Biochem 118: 197-203.

Yan, R, Bienkowski, MJ, Shuck, ME, Miao, H, Tory, MC, Pauley, AM, Brashier, JR, Stratman, NC, Mathews, WR, Buhl, AE, Carter, DB, Tomasselli, AG, Parodi, LA, Heinrikson, RL, Gurney, ME. 1999. Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity. Nature. 402: 533-537.

Yennawar, NH, Yennawar, HP, Farber, GK. 1994. X-ray crystal structure of gamma.-chymotrypsin in hexane. Byochemistry 33:7326-7336.

Zaks, A, Klibanov, A. 1988. The effect of water on enzyme action in organic media. The J Biol Chem. 263(17): 8017-8021

Zale, SE, Klibanov, AM. 1986. Why does ribonuclease irreversible inactivate at high temperatures? Biochem. 25: 5432-5444.

Zarrine-Afsar, A, Larson, SM, Davidson, AR. 2005. The family feud: do proteins with similar structure fold via the same pathway? Curr Opin Struct Biol. 15: 42-49.

Zhu, G, Huang, Que, Qian, M, Tang, Y. 2001. Crystal structure of momorcharin in 80% acetonitrile-water mixtures. Biochim Biophys Acta 1548: 152-158.