

Central modules of the vaccinia virus complement control protein are not in extensive contact

Marina D. KIRKITADZE^{*1}, Colin HENDERSON^{*}, Nicholas C. PRICE[†], Sharon M. KELLY[†], Nicholas P. MULLIN[†], John PARKINSON^{*}, David T. F. DRYDEN[‡] and Paul N. BARLOW^{*}

^{*}Edinburgh Centre for Protein Technology, Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, Scotland, U.K.,

[†]Department of Biological Sciences, University of Stirling, Stirling FK9 4LA, Scotland, U.K., and [‡]Institute of Cell and Molecular Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JR, Scotland, U.K.

The 28.6 kDa vaccinia virus complement control protein (VCP) is an inhibitor of the complement system and has therapeutic potential. It is composed of four domains or modules and is a homologue of complement receptor 1 (CR1) and other mammalian regulators of complement activation. A key aspect to structure–function relationships in these proteins is the extent of intramolecular module–module interactions, since these dictate the overall shape and flexibility of the molecules. A protein fragment (VCP~2,3) encompassing modules 2 and 3 of VCP was over-expressed in *Pichia pastoris*. Ultracentrifugation showed that VCP~2,3 is highly asymmetric with an axial ratio of 5.3:1, which is consistent with an end-to-end arrangement of the two modules. NMR spectroscopy, differential scanning calorimetry, CD and intrinsic tryptophan fluorescence were used

to monitor unfolding of VCP~2,3. Experiments performed over a range of temperatures and concentrations of guanidinium chloride revealed that module 2 unfolds under milder conditions than, and independently of, module 3. Unfolding of module 2 is not associated with extensive changes in amide ¹⁵N and ¹H chemical shifts of module 3, implying that the modules do not form an extensive intermodular interface. Results obtained in this work for VCP~2,3 are compared with those obtained in a study of CR1 modules 15–17 [Kirkkitadze, Krych, Uhrin, Dryden, Smith, Cooper, Wang, Hauhart, Atkinson and Barlow (1999) *Biochemistry* 38, 7019–7031].

Key words: CP modules, unfolding, regulators of complement, short consensus repeat, viral mimicry.

INTRODUCTION

A protein of molecular mass 28.6 kDa, encoded by the viral gene C21L, is present in the medium of cells infected with vaccinia virus [1]. This protein is an inhibitor of both the classical and alternative pathways of complement and serves to defend the virus against complement-mediated attack [2,3]. The structural basis of this biological property is of interest because the inhibition of complement is an important therapeutic goal in numerous clinical situations.

The amino acid sequence of the secreted protein is consistent with the presence of complement protein (CP) modules, also called short consensus repeats. CP modules have been found in more than 50 proteins, but are most prevalent in the six mammalian proteins that regulate complement activation, the regulators of complement activation (RCA) proteins [4]. CP modules each consist of 50–70 amino acid residues and have two disulphide bonds, a virtually invariant tryptophan, and several well-conserved glycines, prolines and hydrophobic residues [5]. They have compact three-dimensional structures with N- and C-termini at opposite ends. Several contiguous modules are required to form each of the binding sites of RCA proteins. Therefore the way in which modules are oriented with respect to their neighbours, and the degree of intermodular flexibility, are of functional relevance.

There are four potential CP modules in the secreted vaccinia virus protein, which together make up its entire sequence, apart from 20 residues at the N-terminus. The solution structure of the C-terminal part of vaccinia virus complement control protein

(VCP), comprising modules 3 and 4 (VCP~3,4), has been solved using NMR [6]. This work showed that these modules are joined in an end-to-end manner, forming an elongated structure with a relatively small intermodular interface (Figure 1, top). ¹⁵N-relaxation studies indicated that the two modules form a fairly rigid unit. This is in contrast to another example of a CP-module pair, modules 15 and 16 of the RCA protein, factor H, where it has proved difficult to define the intermodular orientation [7]. Thus, there is variation in the manner in which adjacent CP modules are arranged with respect to their neighbours. Since all CP modules studied so far have very similar three-dimensional structures (despite low sequence homology) the main challenge in the structural biology of these proteins is to determine the intermodular orientation and flexibility. These are dictated by the nature and the extent of the interface between modules.

The gene for VCP has probably developed from an initial gene-capture event, by which the virus has attained a part, or all, of the host gene and has subsequently adapted the gene for its own purposes. VCP has a functional profile which is more similar to the 30 CP module protein, complement receptor type 1 (CR1) than to any other RCA protein, despite the difference in size between VCP and CR1. The four CP modules of VCP show distinct sequence similarity with modules 1–4, modules 8–11 and modules 15–18 of CR1. These three regions of CR1 are known to recognize and bind to complement proteins C3b and C4b, key components of the complement cascade. Figure 1, bottom, shows a sequence comparison of modules 1, 2 and 3 of VCP, with modules 15, 16 and 17 of CR1. VCP also has high sequence

Abbreviations used: VCP, vaccinia virus complement control protein; VCP~2,3, a fragment of VCP comprising modules 2 and 3; RCA, regulators of complement activation; CP, complement protein; CR1, complement receptor type 1; CR1~16,17 and CR1~15–17, fragments of CR1 comprising modules 16 and 17 and 15–17 respectively; DSC, differential scanning calorimetry; HSQC, heteronuclear single-quantum coherence; GdmCl, guanidinium chloride.

¹ To whom correspondence should be addressed (marina@chem.ed.ac.uk).

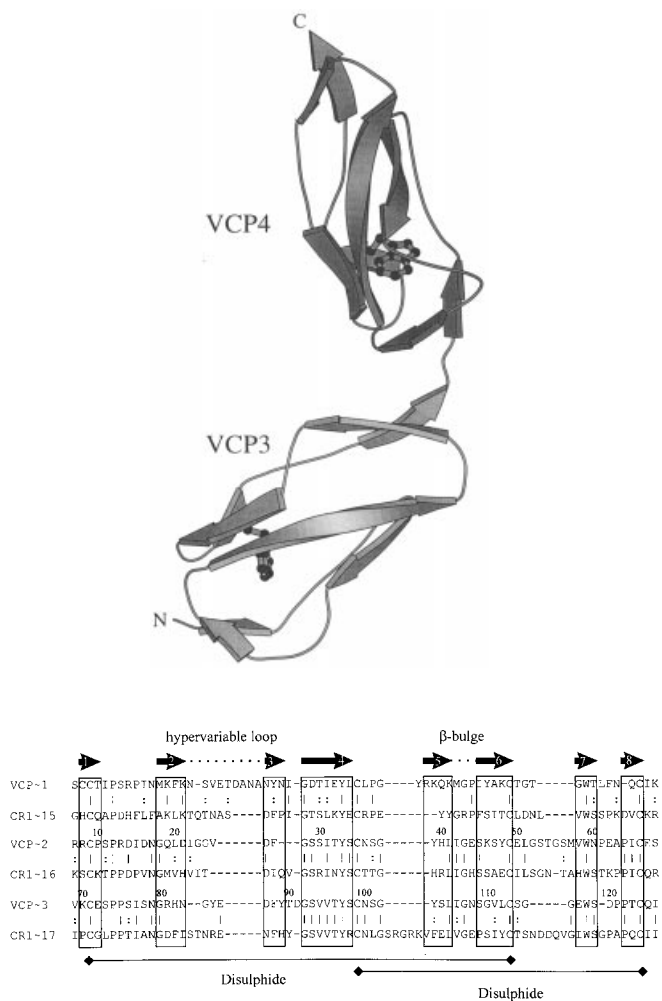


Figure 1 Structure and sequence of CP modules from VCP [6]

(Top panel) Molscript [29] representation of the NMR-derived structure of VCP ~ 3,4 showing the positions of the highly conserved tryptophan residues. (Bottom panel) Sequence alignment for modules 1, 2 and 3 of VCP and modules 15, 16 and 17 of CR1. The bold arrows and the associated boxes indicate the eight short β -strands found in VCP ~ 3 [6]. Strands 2 and 3 are joined by a hypervariable sequence. Solid vertical lines indicate sequence identity, dotted vertical lines indicate conservative substitutions. The two disulphide bonds in each module are indicated. The numerals indicate the amino acid numbering scheme used for VCP ~ 2,3 in this paper.

similarity with modules 1–4 of C4 binding protein chain b and membrane cofactor protein, two other RCA proteins.

In a previous paper [8], we showed how thermodynamic parameters associated with the melting of modules and inter-modular junctions in a C3b-binding site of CR1 (CR1 ~ 15–17) can be interpreted in structural terms. In the present work, a combination of biophysical techniques has been used to study the unfolding of a fragment of VCP and the results are compared with those obtained previously for the equivalent modules of CR1 [8].

EXPERIMENTAL

Purification procedure

The vaccinia virus genome (a gift from Dr. Geoffrey Smith, Department of Pathology, Oxford University, Oxford, U.K.)

was used as a template for the amplification of the coding sequence of modules 2 and 3 of VCP using primers which incorporated *EcoRI* and *NotI* restriction sites at the 5' and 3' ends of the VCP ~ 2,3 coding sequence respectively. The PCR product was ligated into the *Pichia pastoris* expression vector pPICZ α A (Invitrogen, San Diego, CA, U.S.A.). *Escherichia coli* top 10F' cells were transformed with the pPICZ α A/VCP ~ 2,3 construct. Colonies were screened by *EcoRI*–*NotI* digestion of the plasmid DNA (miniprep kit, Qiagen, Hilden, Germany) and plasmid DNA was sequenced using the dye termination method (Amersham Pharmacia, Uppsala, Sweden). A plasmid containing the gene in the correct orientation was purified using the Wizard midiprep method (Promega, Southampton, U.K.) and linearized using the *SaI*I restriction enzyme before transformation of *P. pastoris* mut^s strain, KM71, using electroporation (Bio-Rad Pulser, Hercules, CA, U.S.A.). Transformants were selected by ability to grow on yeast peptone dextran zeocin plates (50 μ g/ml). Subsequently, transformants were screened for protein production by growth in 15 ml of buffered minimal glycerol medium (BMG; prepared as described in the Invitrogen manual) for approx. 24 h at 30 °C, or until an attenuation of between 2 and 6 was reached. Once the desired cell density had been attained, the cells were harvested (5000 *g* for 10 min) and resuspended in 15 ml of buffered minimal methanol medium (BMM; prepared as described in the Invitrogen manual). This cell resuspension was grown for 5 days, with a 1 ml sample of the supernatant being collected each day and analysed by SDS/PAGE. Methanol was added every 24 h to a final concentration of 1% to maintain induction levels. A 1 ml sample of supernatant from a high-expression clone was used for protein sequencing after it had been tested by SDS/PAGE and blotted on to nitrocellulose.

For large-scale expression, the clone was grown overnight in 2 litres of BMG medium at 30 °C in a shaking incubator until a D_{600} of between 2 and 6 was attained. The pellet was then harvested by centrifugation (5000 *g* for 10 min) and resuspended in 400 ml of BMM medium to induce expression. Aliquots were removed and methanol was added at 24 h intervals as before. After 4–5 days of induction, the supernatant was collected (10000 *g* for 1 h, the pellet was discarded) and concentrated to ~ 50 ml using an Amicon concentrator (3 kDa cut-off) before being centrifuged at 20000 *g* for 1 h, retaining only the supernatant for further analysis. The supernatant was applied to a PD-10 (Amersham Pharmacia) gel-filtration column, previously equilibrated with 5 mM sodium acetate, pH 4.0, and eluted in a volume of 3.5 ml. The eluate was filtered through a 0.2 μ m membrane and applied to a Mono-S cation-exchange column (Amersham Pharmacia) equilibrated in 5 mM sodium acetate, pH 4.0. A linear gradient of 0–1 M NaCl (20 min) was applied and VCP ~ 2,3 was eluted at 0.6 M NaCl. The absorbance was monitored at 280 nm. All fractions corresponding to absorbance peaks were analysed by SDS/PAGE for the presence of VCP ~ 2-3. For MS analysis, the fractions containing pure VCP ~ 2,3 were desalted using a C_4 reverse-phase column, eluted with an acetonitrile gradient, freeze-dried, dissolved in 50% acetonitrile in aqueous solution containing 0.1% formic acid and analysed by electrospray-ionization MS (Micromass, Manchester, U.K.). The mass obtained was 13582 Da, which is within one mass unit of the expected mass of the VCP ~ 2,3 sequence plus one EAEFIK at the N-terminus, which is a cloning artefact. For NMR experiments, freeze-dried VCP ~ 2,3 was dissolved in 550 μ l of 10 mM sodium phosphate, pH 6.0, and 50 μ l of $^2\text{H}_2\text{O}$.

To prepare the ^{15}N -labelled sample, the nitrogen source for both BMG and BMM media, normally yeast nitrogen base (YNB), was replaced with YNB which had no amino acids present (Difco). ^{15}N -Ammonium sulphate was added to a con-

centration of 0.2% and the growth carried out as described above.

For all experiments a buffer containing 10 mM sodium phosphate, pH 6.0, was used.

Analytical ultracentrifugation

Analytical ultracentrifugation experiments were performed using Beckman Optima XL-A and XL-I analytical centrifuges with absorption and Rayleigh interference optics. Both the centrifuges had full on-line computer data capture and analysis facilities (Beckman, Palo Alto, CA, U.S.A.). Sedimentation velocity experiments were performed at 4 °C at a rotor speed of 55000 rev./min using a 12 mm optical path-length cell containing 400 μ l of sample. Sedimentation coefficients were evaluated from the absorption traces using the Svedberg algorithm as described by Philo [9]. The axial ratio a/b was evaluated from the Perrin shape function, P , for ellipsoids of revolution [10] using the routine ELLIPS1 [11].

Differential scanning calorimetry (DSC)

DSC studies were conducted on an MC-2 differential scanning calorimeter (Microcal, Northampton, MA, U.S.A.) with the assistance of Professor Alan Cooper and Margaret Nutley (University of Glasgow, Glasgow, Scotland, U.K.). The cell volume was 1.5 ml, the rate of heating was 1 °C/min, and excess pressure was kept equal to 80 bar (1 bar = 10^5 Pa). The partial molar heat capacity and melting curve were analysed using standard procedures [12]. The data were processed using the software ORIGIN.2 (Microcal). The protein concentration was 0.7–3 mg/ml as determined by measurements of absorbance at 278 nm according to a calculated [13] absorption coefficient of 1.2 for a 1 mg/ml solution of VCP~2,3 in a path-length of 1 cm.

CD

CD measurements were performed using a Jasco-600 spectropolarimeter (Japan Spectroscopic Company, Tokyo, Japan) with a cylindrical quartz cell of path-length 0.02 cm. The protein concentration was 0.2 mg/ml. All measurements were recorded at 25 °C except where stated. In order to obtain plots of percentage total change versus temperature, the maximum ellipticity and reading at ± 1.0 nm of the maximum were averaged to smooth experimental errors in the curves.

NMR

NMR measurements were performed with the assistance of Dr. Dusan Uhrin (University of Edinburgh, Edinburgh, Scotland, U.K.) on a Varian (Palo Alto, CA, U.S.A.) Inova NMR spectrometer operating at 600 MHz (proton frequency) in a 5 mm triple-resonance pulsed-field gradient probe. The ^{15}N - ^1H heteronuclear single-quantum coherence (HSQC) spectra of 0.3–0.6 mM samples of recombinant ^{15}N -labelled protein were acquired for 50 min each, at a range of temperatures from 25 to 55 °C, using the composite-pulse Watergate for water suppression [14]. HSQC spectra of VCP~2,3 in the presence of guanidinium chloride (GdmCl) were collected at 25 °C. Data were processed and plotted using AZARA (Wayne Boucher, Department of Chemistry, University of Cambridge, U.K.).

Intrinsic tryptophan fluorescence measurements

Intrinsic tryptophan fluorescence measurements were performed on a Perkin-Elmer LS-50B spectrofluorimeter in a 100 μ l cuvette

of 0.3 cm path-length at 25 °C. The excitation wavelength was 295 nm and the emission spectra were recorded over 300–500 nm. The spectral bandwidth was 10 nm and protein concentration was 0.3 mg/ml. The change in fluorescence intensity as a function of denaturant was converted into fraction of protein unfolded by linear fitting in regions at low and high denaturant concentrations and then analysed by GRAFIT (Erithacus Software, Staines, U.K.) with equations [15] for one or two independent two-state unfolding transitions.

Molecular modelling

Models of the tertiary structure of VCP~2, CR1~15 and CR1~16 were constructed with the program MODELLER [16] using homology to the known single and double CP-module structures. Molecular dimensions of individual modules were obtained by first aligning the modules along their principal moments of inertia using the program 'Inertia' (available from Dr A. R. C. Raine, Department of Chemistry, University of Cambridge, Cambridge, U.K.). To avoid the problem of extended side-chains artificially increasing the apparent dimensions of the modules, molecular dimensions were defined as the distances along each of the three principal axes from the centre of mass of the structure, within which 95% of the heavy atoms were found. Axial ratios of the modules were then calculated by dividing the longest dimension by the mean of the two shortest dimensions. 'Core' atoms within the structures were defined as those atoms with no accessible surface as calculated by NACCESS (available from the authors S. J. Hubbard and J. M. Thornton, Department of Biochemistry and Molecular Biology, University College London, London, U.K.). Hydrogen bonds were calculated using the program HBPLUS [17]. Volume and accessible surface area were calculated using GRASP [18].

RESULTS

Analytical ultracentrifugation studies and the shape of VCP~2,3

The average value of the Svedberg coefficient ($s_{20,w}^0$) evaluated

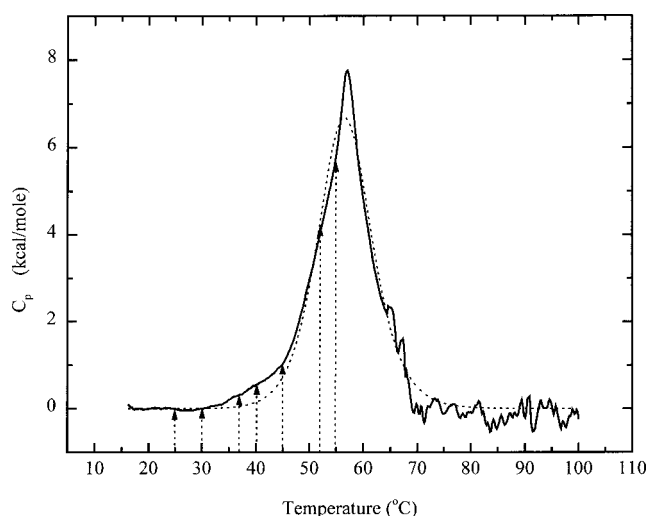


Figure 2 Differential scanning calorimetry profile of VCP~2,3

Experimental data after base-line subtraction are shown by the solid line. The broken line indicates the computer-fitted data. Arrows indicate the temperatures at which NMR data were collected (see Figure 3).

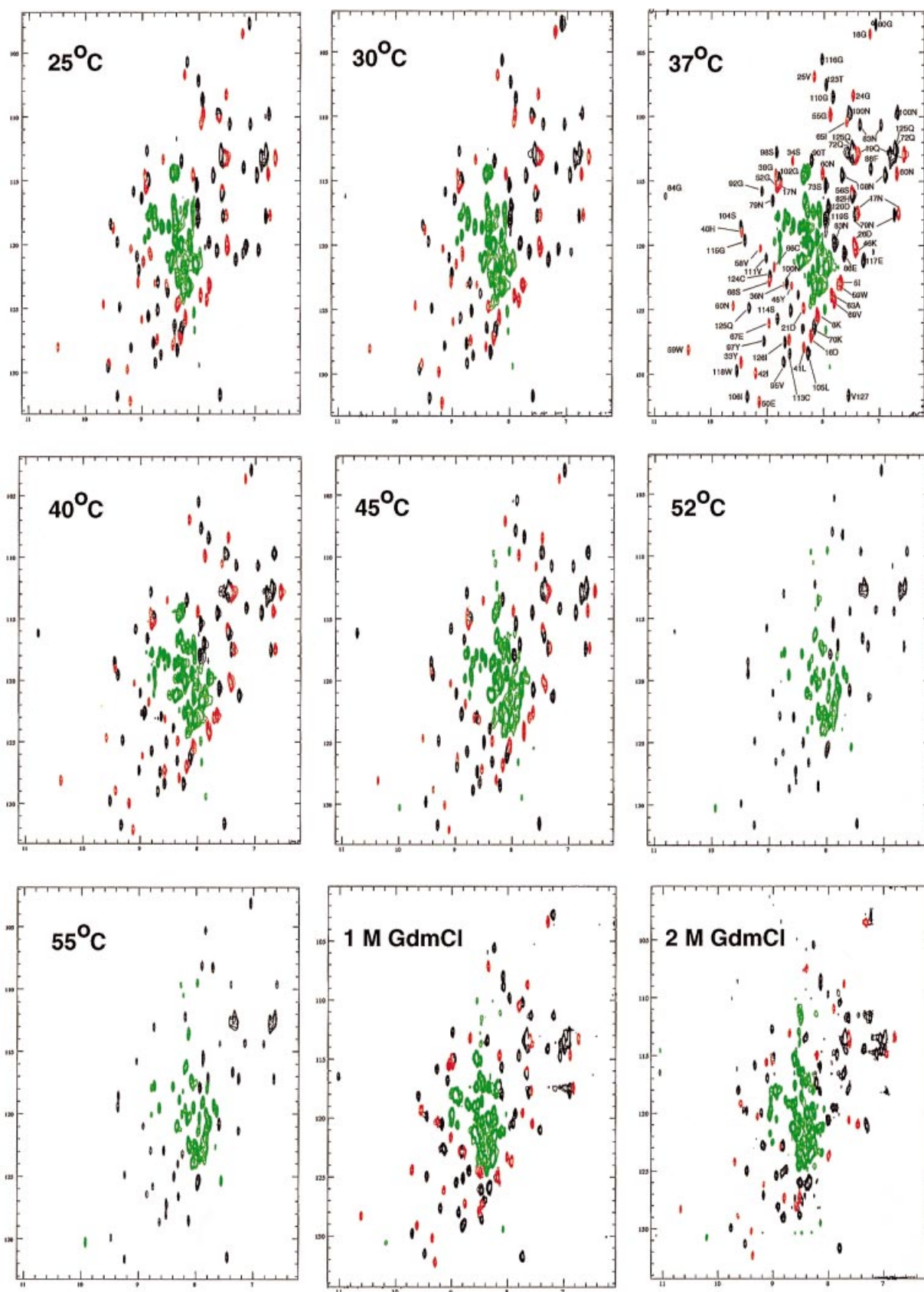


Figure 3 Series of ^1H ^{15}N -HSQC spectra collected on VCP ~ 2,3

Cross-peaks assigned to module 2 and module 3 are in red and black respectively. Cross-peaks where module assignment is difficult, because they are in a crowded position or because they correspond to random-coil positions, are in green. Spectra were recorded at various temperatures as indicated, and (bottom right) in the presence of 1 and 2 M GdmCl at 25 °C. For the spectrum collected at 37 °C, sequence-specific assignments are shown for well-dispersed cross-peaks.

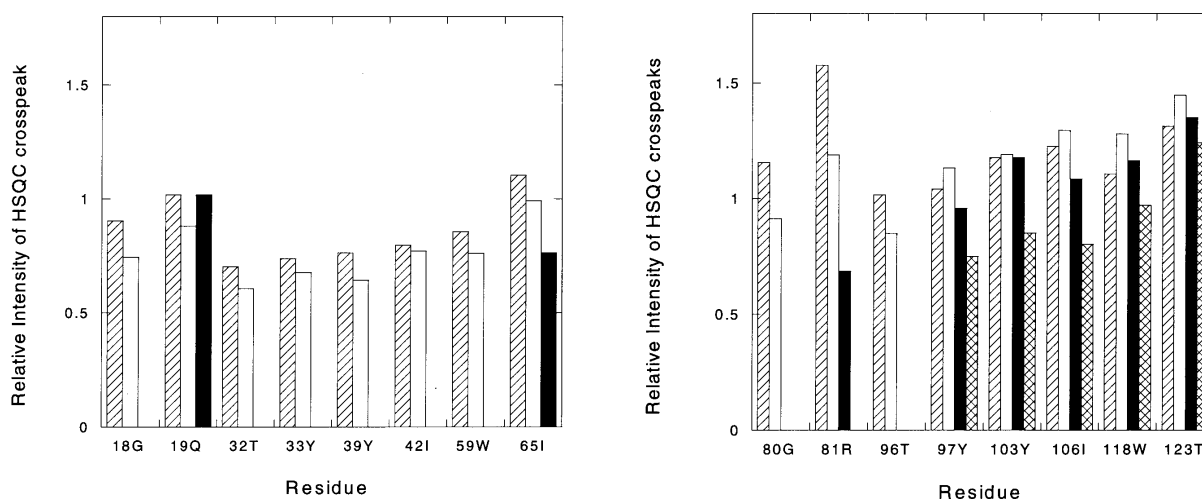


Figure 4 Slowly exchanging protons in VCP ~ 2,3

Intensity of HSQC cross-peaks (arbitrary units) after a freeze-dried sample of ^{15}N -labelled VCP ~ 2,3 had been dissolved in $^2\text{H}_2\text{O}$ for 13 min (hatched bars), for 30 min (white bars), for 60 min (black bars), and for 120 min (cross-hatched bars). Panels (a) and (b) show values of intensity for equivalent residues (see Figure 1b) of modules 2 and 3 respectively.

from sedimentation velocity experiments (over a protein concentration range of 1–5 mg/ml) was $s_{20,w}^0 = 1.44 \pm 0.01$ S. Corresponding values for the axial ratio for VCP ~ 2,3 calculated assuming three different values of hydration ($\delta = 0.2, 0.35, 0.5$) were: 5.6:1, 5.3:1 and 5.0:1. Therefore it may be concluded that VCP ~ 2,3 under native conditions is clearly an asymmetric molecule with an axial ratio ~ 5.3:1.

Temperature-induced unfolding of VCP ~ 2,3

DSC studies

The thermal unfolding of VCP ~ 2,3 has been studied using DSC. A plot of partial heat capacity against temperature shows a transition with a midpoint at 57 °C (Figure 2). The value for the calorimetric enthalpy, 86 kcal/mol (1 cal \equiv 4.184 J), is consistent with the value obtained for other globular proteins of a similar size [12].

A distinct shoulder on the calorimetric profile, and a value > 1 for the ratio of calorimetric and van't Hoff enthalpies ($\Delta H^{\text{cal}}/\Delta H^{\text{eff}}$), both indicate that the protein does not represent a single co-operative system, but rather that melting of the various components occurs at similar, but not identical, temperatures [19]. In addition to the major transition at 57 °C, there is a minor transition in the 30–40 °C range. In repeated scans, this was present either as a discrete peak or as a shoulder on the edge of the main feature. By using the calorimetry profile as a guide, NMR spectroscopy can be performed on VCP ~ 2,3 at different temperatures to ascertain the order in which components melt.

NMR studies

Two-dimensional $^1\text{H}, ^{15}\text{N}$ -HSQC spectra correlate amide proton and nitrogen chemical shifts. Thus in a $^1\text{H}, ^{15}\text{N}$ -HSQC spectrum of a protein, a cross-peak may be expected for each non-proline amino acid and for some side-chains. Formation of secondary and tertiary structure results in dispersion of HSQC cross-peaks over a wide range of chemical shifts. On the other hand, amide protons within a random-coil polypeptide have a very narrow range of proton chemical shifts and tend to overlap. Conform-

ational exchange results in broadening or doubling of peaks. Thus, an HSQC spectrum provides a reliable method of assessing whether a protein or protein domain is compactly folded or not.

In the case of ^{15}N -labelled VCP ~ 2,3, most of the cross-peaks in a $^1\text{H}, ^{15}\text{N}$ -HSQC spectrum collected at 37 °C were assigned to specific amino acid residues (Figure 3) using a set of standard 2- and 3-dimensional NMR techniques. This allows the monitoring of individual amino acids at different temperatures. Figure 3 (top-left panel) shows the $^1\text{H}, ^{15}\text{N}$ -HSQC spectrum of VCP ~ 2,3 collected at 25 °C on a 1 mM sample and displaying approx. 120 well-dispersed, sharp, cross-peaks, consistent with a fully-folded, monodisperse protein. Figure 3 also shows $^1\text{H}, ^{15}\text{N}$ -HSQC spectra collected at 30, 37 (annotated with assignments), 40, 45, 52 and 55 °C respectively. Many of the well-dispersed peaks do not show extensive changes in chemical shift with increasing temperature and can be identified readily (Figure 3). The small deviation in the position of peaks which occurs between 25 and 40 °C indicates that no significant unfolding of modules has occurred over this temperature range. Between 40 and 45 °C, the cross-peaks assigned to module 2 become lower in intensity and new cross-peaks consistent with the presence of random-coil structure appear (e.g. Trp side-chain at $^1\text{H}_\beta = 10$ p.p.m. for H^β , $^{15}\text{N}_\beta = 130$ p.p.m. for N^β). The disappearance or splitting of cross-peaks observed in the range 45–55 °C (Figure 3) corresponds mainly to the loss of tertiary structure in module 2. This may be concluded, since at 52 °C most of the dispersed peaks assigned to module 2 (1–68) have moved and are not identifiable or have disappeared. In contrast, many of the cross-peaks corresponding to module 3 (69–127) are still recognizable at 55 °C. Thus, signals arising from module 2 fade rapidly with increasing temperature, whereas the majority of peaks of module 3 remain well dispersed. These data indicate that module 2 of the VCP ~ 2,3 fragment is less stable to temperature than module 3.

Figure 4 shows the intensity of selected HSQC cross-peaks from equivalent residues in modules 2 and 3 after incubation of VCP ~ 2,3 in 99.8% $^2\text{H}_2\text{O}$ at 37 °C for various periods of time between 13 and 120 min. Under these conditions, amide protons may be replaced by deuterons through chemical exchange. This process appears to occur more rapidly in module 2 than in module 3. There are very few cross-peaks for module 2 detectable

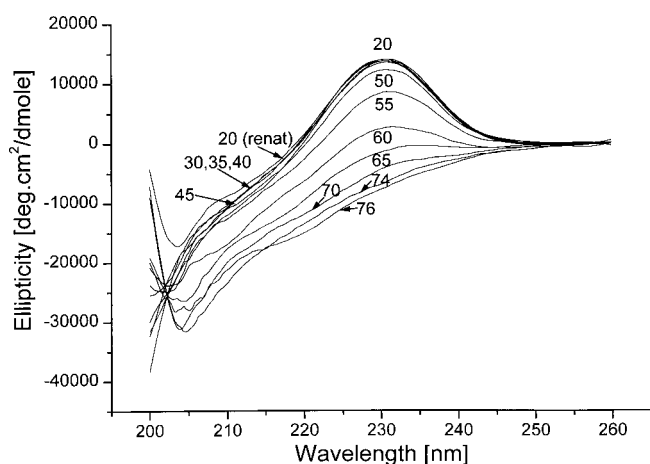


Figure 5 Temperature-induced unfolding followed by CD

A series of far-UV CD spectra for VCP~2,3 were collected over a range of temperatures, as indicated by the numerals on the Figure. A spectrum was collected at 20 °C after cooling of the sample which resulted in protein renaturation (renat).

after 60 min, in contrast to module 3. In the case of both modules, those amides which exchange more slowly correspond to the predicted β -sheet structure for CP modules. These data indicate that at 37 °C, module 2 is less compactly folded than module 3.

CD studies

Far-UV CD spectra of proteins are dominated by peptide-bond dichroism and are sensitive to the extent and nature of secondary structure. CD spectra of VCP~2,3 revealed an unusual, strong positive signal in the far-UV region. Figure 5 shows CD spectra of VCP~2-3 obtained over a range of temperatures, 20–75 °C.

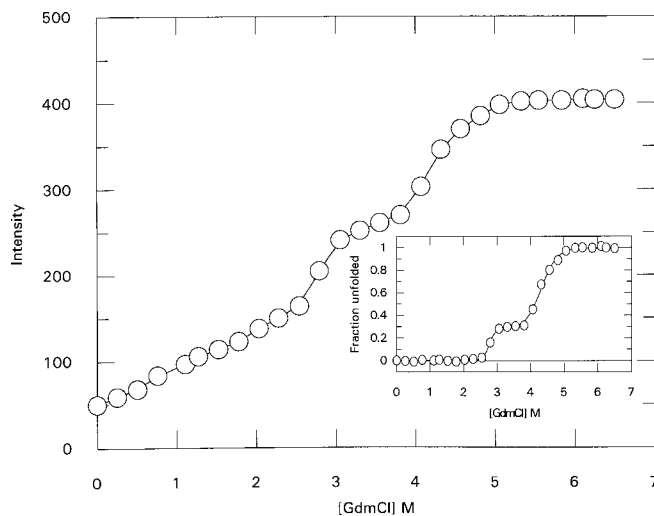


Figure 6 GdmCl-induced unfolding monitored by intrinsic tryptophan fluorescence

Intensity of fluorescence at 340 nm as a function of GdmCl concentration. In order to extract thermodynamic parameters from the two sharp transitions, the data were fitted to two independent transitions (inset) as described in the text.

Comparatively small changes (6% loss of the characteristic positive peak at 230 nm) are observed below 40 °C. An approx. 30% loss of the CD signal takes place at 55 °C and a 55% loss occurs at 60 °C, whereas at 70 °C the peak has disappeared and the CD signal indicates that the structure is that of a random coil. This behaviour is in good agreement with the DSC data. On cooling to 20 °C, the characteristic signal at 230 nm is regained, indicating that the thermal unfolding is reversible.

GdmCl-induced unfolding of VCP~2,3

Fluorescence studies

GdmCl-induced equilibrium unfolding of VCP~2,3 has been monitored by intrinsic tryptophan fluorescence. Each module has a conserved Trp and there are no Trp residues elsewhere in the sequence (Figure 1, top). On the basis of homology with experimentally derived structures, the conserved Trp side-chains can be expected to occupy positions within the hydrophobic cores of the modules, close to the disulphide bond and the junction with the next module. The increase in the intensity of tryptophan fluorescence upon denaturation of VCP~2,3 at 340 nm was about 10-fold (Figure 6). An initial increase in intensity, corresponding to 30% of the total change, occurs between 0 and 2.5 M GdmCl. This is followed by two major structural transitions with midpoints at 2.75 M and 4.5 M GdmCl; these corresponded to free energy changes of $81.6 \pm 22.8 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $45.3 \pm 3.2 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, and transition slopes of $29.6 \pm 8.2 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $10.5 \pm 3.2 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ respectively. To help analyse these titrations, NMR experiments have been performed on VCP~2,3 at different concentrations of GdmCl.

NMR studies of VCP~2,3 in GdmCl

$^1\text{H}, ^{15}\text{N}$ -HSQC spectra were collected and analysed to assess the stability of VCP~2,3 modules in the presence of 1 and 2 M GdmCl. In 1 M GdmCl, strong, sharp cross-peaks corresponding to both modules 2 and 3 are present (Figure 3) and there is only a slight indication of random-coil structure. The initial increase in relative fluorescence intensity (about 12% of the total change) is therefore probably not entirely due to module unfolding, and could well correspond to unfolding of the intermodule interface. At 2 M GdmCl, the intensity of cross-peaks for both modules 2 and 3 decreases, but the loss of intensity is more marked for module 2 than for module 3 (Figure 3). This implies that at 2 M GdmCl, module 2 is unfolding to a greater extent than module 3 and that module 2 is partially unfolded prior to the mid-point of the first major transition monitored by fluorescence.

Modelling studies

Given the good conservation of 'scaffold' residues among CP modules of VCP and CR1, it is possible to build reliable structural models of individual CP modules on the basis of homology. In order to rationalize the data obtained for the unfolding of VCP modules 2 and 3, and to compare them with the results of previous studies of the homologous modules 16 and 17 from CR1, a modelling exercise was undertaken.

The three unknown (VCP~2, CR1~16 and CR1~17) CP-module structures were modelled by homology with known CP-module structures, including that of VCP~3 [6]. In each case, five modelled structures were generated. The values for the axial ratios of the four single modules were: 2.1:1 for CR1~16, 2.2:1 for CR1~17, 2.0:1 for VCP~2 and 1.9:1 for VCP~3. As expected, the conserved Trp residue was found to be buried

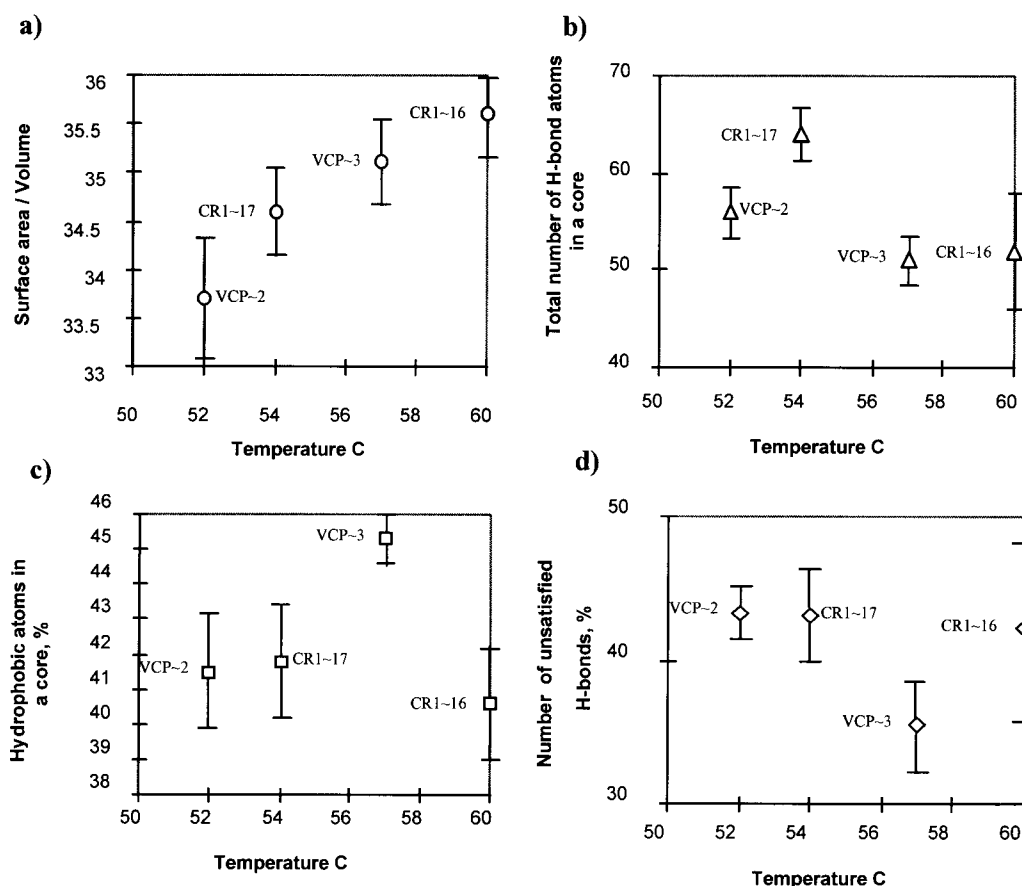


Figure 7 Analysis of features in CP-module structures that might contribute to thermal stability

The calculated data are plotted against measured melting temperatures of specific modules. Error bars are S.D.s based on analyses of five modelled or experimentally derived structures. (a) Surface area/volume ratio. (b) Total number of atoms with the potential to form hydrogen bonds in the core. (c) Percentage of hydrophobic atoms in the core. (d) Percentage of unsatisfied hydrogen-bonds.

within the core of each of the model structures. A number of structural analyses of these models and of five experimentally determined structures of VCP ~ 3 [6] were performed (see Figure 7). VCP ~ 3 was found to have the lowest number of unsatisfied core hydrogen-bonds and the greatest number of hydrophobic atoms buried in the core of its structure. VCP ~ 2 was found to have a lower surface area to volume ratio than CR1 ~ 16 and VCP ~ 3, with CR1 ~ 17 lying somewhere in the middle.

DISCUSSION

The primary sequence of VCP is consistent with the presence of four CP modules joined by short linkers. The NMR solution structure of the C-terminal half of VCP (modules 3 and 4) confirmed that two of the CP-like sequences form discrete compactly folded modules, and that these are joined end-to-end, forming a fairly rigid rod-like structure. Residues forming the intermodular interface are not highly conserved and it is therefore difficult to extrapolate to other intermodular interfaces in VCP or elsewhere in the RCA family. A recent study [8] showed that a combination of hydrodynamic and thermodynamic methods, with the guidance of assigned NMR spectra, could address the issue of how closely CP modules associate with one another. This study was performed on modules 15–17 of CR1. In the current study, a similar approach has been taken to the central

two modules of VCP, modules 2 and 3, which are very similar in sequence to modules 16 and 17 of CR1 respectively.

The VCP fragment, VCP ~ 2,3, was successfully expressed in *P. pastoris*, and the authenticity of the purified material was confirmed by electrospray ionization MS. This confirmed the presence of six extra residues at the N-terminus (EAEFIK), which were anticipated cloning artefacts [20]. NMR studies showed that the recombinant protein was fully folded and monodisperse at concentrations up to 10 mg/ml. The HSQC spectra were assigned using standard techniques. The spectrum contains two similar sets of roughly 60 cross-peaks, consistent with the presence of two properly folded, structurally related modules. As temperature is increased the cross-peaks ascribed to module 2 lose their chemical-shift dispersion due to module unfolding, but the chemical shifts of module 3 do not change significantly up until 52 °C. Furthermore, comparison of chemical shifts with VCP ~ 3,4 indicated that module 3 of VCP ~ 2,3 had folded in a very similar way to module 3 of VCP ~ 3,4, i.e. it folds independently of modules 2 and 4. These results show that modules 2 and 3, like modules 3 and 4, do not interact extensively with each other. Nevertheless, it has been shown that the four CP modules of VCP do act in a concerted fashion to bind ligand. Therefore the problem of interfaces remain a key issue in understanding structure–function relationships of the protein.

The axial ratio of approx. 5.3:1 derived from ultracentrifugation shows that, under native conditions, VCP~2,3 is a highly asymmetric molecule. Since individual modules are known from NMR studies to have axial ratios of approx. 2:1, and the modelled structures of VCP~2 and VCP~3 each have an axial ratio of ~2:1, this result suggests that the prolate ellipsoid model used to calculate axial ratios from the hydrodynamic data is not an optimal approximation. However, the results do appear to be consistent with the two CP modules being arranged in an end-to-end fashion, as found in VCP~3,4 [6]. The axial ratio of VCP~2,3 is also similar to that obtained for the very close homologue CR1~16,17 (axial ratio 5:1) [21]. It was shown earlier that other molecules of complement also adopt elongated structures. For example the frictional ratio of a recombinant, soluble form of CR2 (containing 15 CP modules) indicates that it is a long rod-shaped molecule [22]. Factor H was also revealed to have a flexible extended structure [23].

It has been shown previously that Trp contributes to the CD signal in both the far- and near-UV wavelength regions [24]. Assignment of the contribution of the Trp residues to the CD spectrum of human carbonic anhydrase II made by Freskgård et al. [25] shows that Trp residues, all positioned in various β -strands, make a significant, positive contribution in the 220–240 nm region. Members of the complement family, such as CR1~15–17 [8], the CR1~1–3 fragment [26], DAF~3,4 and the N-terminal fragment of the metabotropic γ -aminobutyric acid receptor [27] (which is composed of a tandem pair of complement protein modules), also show positive ellipticity in the far-UV region. This appears to be a signature of folded CP modules and is likely to reflect certain conserved features within their structures which typically comprise eight short β -strands connected by loops. The nature of the CD profile may be related to the invariant core Trp residues which are located on β -strands and lie close to one of the two totally conserved disulphide bridges in each module.

As the temperature is increased from 20 to 40 °C, the value of ellipticity for VCP~2,3 at 230 nm (Figure 5) decreases by about 6%. A minor transition on the calorimetry profile is detected between 32 and 40 °C, giving rise to a shoulder on the leading edge of the major 57 °C transition (Figure 2) (or a small, distinct peak at a higher protein concentration). Figure 3 shows that many of the ^1H - ^{15}N cross-peaks do not show significant changes in chemical shift between 20 and 40 °C, which indicates that both modules of VCP~2,3 remain fully folded over this temperature range. A minor increase (12%) of the relative fluorescence intensity was observed over the range 0–1 M GdmCl for VCP~2,3, although NMR indicates very little unfolding of modules. These small adjustments of VCP~2–3 conformation, observed by a combination of methods under mild conditions, could correspond to melting of the intermodular junction. If so, melting of the junction between modules 2 and 3 of VCP takes place at a lower temperature (or at a lower GdmCl concentration), and with an apparently smaller enthalpy value, than the junction between modules 16 and 17 of CR1 [8]. Modules 2 and 3 of VCP are homologous with modules 16 and 17 of CR1 (Figure 1b). Thus, according to this interpretation, the VCP~2,3 intermodular junction appears to be less stable and extensive than that between CR1~16–17.

The gradual increase below 2.5 M GdmCl in the fluorescence intensity may correspond both to melting of the junction and to partial melting of module 2. The subsequent sharp transition at 2.75 M GdmCl could correspond to partial unfolding of module 3, which shows some signs of unfolding even at 2 M GdmCl, and further unfolding of module 2. In the case of CR1~15–17, it was possible to evoke a molten form of the modules at 5 M GdmCl

lacking in tertiary structure but with tryptophan buried in a hydrophobic cluster; this became further unfolded in a second transition at 5.4 M GdmCl. Similarly, for VCP~2,3 there is a second transition at 4.5 M GdmCl and this could correspond to the further unfolding of an analogous, molten form of modules 2 and 3.

The analysis of VCP~2,3 indicates that module 2 unfolds more readily than, and independently of, module 3. The stability of modules 2 and 3 of VCP may be compared with the homologous modules 16 and 17 of CR1. Module 2 of VCP~2,3 melts at 52 °C, whereas module 16 of CR1~15–17 undergoes melting at 60 °C [8]. Module 3 of VCP melts at 57 °C, whereas module 17 of CR1~15–17 does so in the range 50–57 °C. Thus module 2 of VCP appears to be less stable than the homologous module 16 of CR1, but module 3 of VCP is more stable than the homologous module 17 of CR1. Consequently, the central modules of VCP melt independently of one another and in a different order to the homologous two modules of CR1.

In an attempt to rationalize the experimental data, we performed an analysis of a number of factors thought to be important in determining module stability. However, it is clear that there are problems relating from the uncertainty of the magnitudes and definitions of these various components (Figure 7). There is no positive or negative correlation between melting temperature and the size of the hydrophobic core, the number of unsatisfied hydrogen bonds or the total number of hydrogen bonds. There does appear to be a weak positive correlation between the surface area/volume ratio and melting temperature, although more data points would be required to confirm this. Such a correlation would be expected, as proteins with a small change in surface area on denaturation are more stable than those which give a large change in surface area on denaturation [28].

Conclusions

The approach of combining several biophysical techniques to study unfolding of CP modules and intermodular junctions used previously for CR1 [8] has been applied to VCP~2,3 expressed in *P. pastoris*. The protein fragment has been shown to contain two compactly folded CP modules joined in an end-to-end fashion as they are in CR1~15–17 and in the NMR-determined structure of VCP~3,4. The two modules melt independently of one another and in a different order to the homologous modules of CR1. The 2–3 intermodular junction appears to be less stable, and presumably less extensive, than the 16–17 junction of CR1.

We thank Margaret Nutley and Professor Alan Cooper (University of Glasgow) for their kind and expert assistance with the DSC studies. We are also very grateful to Professor Stephen E. Harding and Dr. Kornelia Jumel (University of Nottingham) for their help with recording and analysing the ultracentrifugation data. M. D.K. is funded by the Human Frontiers Science program. The Edinburgh Centre for Protein Technology is funded by the Biotechnology and Biological Sciences Research Council (BBSRC). D.T.F.D. holds a University Research Fellowship funded by the Royal Society of Great Britain. The CD facility in Stirling University is funded by the BBSRC.

REFERENCES

- McCrae, M. A. and Pennington, T. H. (1978) *Virology* **95**, 828–834
- Kotwal, G. J., Isaacs, S. N., McKenzie, R., Frank, M. M. and Moss, B. (1990) *Science* **250**, 827–830
- Isaacs, S. N., Wolfe, E. J., Payne, L. G. and Moss, B. (1992) *J. Infect. Dis.* **66**, 7217–7224
- Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T. and Tack, B. F. (1986) *Immunology Today* **7**, 230–234
- Reid, K. B. M. and Day, A. J. (1988) *Immunology Today* **10**, 177–180
- Wiles, A. P., Shaw, G., Bright, J., Perczel, A., Campbell, I. D. and Barlow, P. N. (1997) *J. Mol. Biol.* **272**, 253–265
- Barlow, P. N., Steinkasserer, A., Norman, D. G., Kieffer, B., Wiles, A. P., Sim, R. B. and Campbell, I. D. (1993) *J. Mol. Biol.* **232**, 268–284

- 8 Kirkitadze, M. D., Krych, M., Uhrin, D., Dryden, D. T. F., Smith, B. O., Cooper, A., Wang, X., Hauhart, R., Atkinson, J. P. and Barlow, P. N. (1999) *Biochemistry* **38**, 7019–7031
- 9 Philo, J. S. (1997) *Biophys. J.* **72**, 435–444
- 10 Van Holde, K. E. (1985) *Physical Chemistry*, Prentice Hall, Englewood Cliffs, NJ
- 11 Harding, S. E. and Cölfen, H. (1995) *Anal. Biochem.* **228**, 131–142
- 12 Privalov, P. L. and Potekhin, S. A. (1986) *Methods Enzymol.* **131**, 4–51
- 13 Sober, H. A. (1970) *Handbook of Biochemistry*, 2nd edn., pp. B75–B76, CRC Press Inc., Boca Raton, FL
- 14 Mori, S., Abeygunawardana C., Johnson M. O. and van Zijl, P. C. M. (1995) *J. Magn. Reson. B* **108**, 94–98
- 15 Hamaguchi, K. (1992) *The Protein Molecule: Conformation Stability and Folding*, pp. 115–160, Japan Scientific Societies, Tokyo/Springer Verlag, Berlin
- 16 Sali, A. and Blundell, T. L. (1993) *J. Mol. Biol.* **234**, 779–815
- 17 McDonald, I. K. and Thornton, J. M. (1994) *J. Mol. Biol.* **238**, 777–793
- 18 Nicholls, A., Bharadwaj, R. and Honig, B. (1993) *Biophys. J.* **64**, A166
- 19 Privalov, P. L. (1982) *Adv. Protein Chem.* **35**, 1–104
- 20 Denton, H., Smith, M., Uhrin, D., Uhrinova, S., Barlow, P. N. and Sawyer, L. (1998) *Protein Expression Purif.* **14**, 97–103
- 21 Kirkitadze, M. D., Jumel, K., Harding, S. E., Dryden, D. T. F., Krych, M., Atkinson, J. P. and Barlow, P. N. (1999) *Prog. Colloid Polymer Sci.* **113**, 164–167
- 22 Moore, M. D., DiScipio R. G., Cooper N. R. and Nemarrow, G. R. (1989) *J. Biol.Chem.* **264**, 20576–20582
- 23 DiScippio, R. G. (1992) *J. Immunol.* **149**, 2592–2599
- 24 Honig, B. and Yang, A.-S. (1995) *Adv. Protein Chem.* **46**, 27–58
- 25 Freskgård P.-O., Mårtensson L.-G., Jonasson P., Jonsson B. H. and Carlsson, U. (1994) *Biochemistry*, **33**, 14281–14288
- 26 Clark, N. S., Dodd, I., Mossakowska, D. E., Smith, R. A. G. and Gore, M. G. (1996) *Protein Eng.* **9**, 877–884
- 27 Hawrot, E., Xiao, Y., Shi, Q.-L., Norman, D., Kirkitadze, M. and Barlow, P. N. (1998) *FEBS Lett.* **432**, 103–108
- 28 Fersht, A. (1999) *Structure and Mechanism in Protein Science*, pp. 508–539, W. H. Freeman and Co., New York
- 29 Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950

Received 5 July 1999/6 August 1999; accepted 9 September 1999