Independently Melting Modules and Highly Structured Intermodular Junctions within Complement Receptor Type 1†

Marina D. Kirkitadze,‡ Malgorzata Krych,§ Dusan Uhrin,‡ David T. F. Dryden,|| Brian O. Smith,‡ Alan Cooper,⊥ Xuefeng Wang,§ Richard Hauhart,§ John P. Atkinson,§ and Paul N. Barlow*,‡

The Edinburgh Centre for Protein Technology, Joseph Black Building, West Mains Road, Edinburgh EH9 3JJ, Scotland, Department of Internal Medicine, Division of Rheumatology, Washington University School of Medicine, 660 South Euclid, Box 8045, Saint Louis, Missouri 63110, Institute of Cell and Molecular Biology, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JJ, Scotland, and Chemistry Department, Glasgow University, Glasgow, G12 8QQ, Scotland

Received October 14, 1998; Revised Manuscript Received February 22, 1999

ABSTRACT: A segment of complement receptor type 1 (CR1) corresponding to modules 15−17 was overexpressed as a functionally active recombinant protein with N-glycosylation sites ablated by mutagenesis (referred to as CR1∼15−17$^-$). A protein consisting of modules 15 and 16 and another corresponding to module 16 were also overexpressed. Comparison of heteronuclear nuclear magnetic resonance (NMR) spectra for the single, double, and triple module fragments indicated that module 16 makes more extensive contacts with module 15 than with module 17. A combination of NMR, differential scanning calorimetry, circular dichroism, and tryptophan-derived fluorescence indicated a complex unfolding pathway for CR1∼15−17$^-$. As temperature or denaturant concentration was increased, the 16−17 junction appeared to melt first, followed by the 15−16 junction, and module 17 itself; finally, modules 15 and 16 became denatured. Modules 15 and 16 adopted an intermediate state prior to total denaturation. These results are compared with a previously published study [Clark, N. S., Dodd, I, Mossakowska, D. E., Smith, R. A. G., and Gore, M. G. (1996) *Protein Eng. 9*, 877−884] on a fragment consisting of the N-terminal three CR1 modules which appeared to melt as a single unit.

Most extracellular proteins are mosaic in nature, being built up from protein modules (*1*). The same is true of the extracellular portions of many cell-surface receptors. Examples include proteins of the immune system, the complement and clotting cascades, and the extracellular matrix. Numerous intracellular proteins also have a modular composition, and examples include many of the proteins involved in signaling (*2*). Nature has utilized a relatively small number of different module-types, possibly fewer than 100 (*3*), to build thousands of functionally distinct proteins. The three-dimensional structures of many examples of individual modules are now available (reviewed in ref *3*). Each module-type is characterized by a consensus sequence which normally defines a common three-dimensional molecular framework. Loops and exposed side chains are not conserved among modules of the same category, and these enable a diversity of binding properties. Further diversity may result from the combination of several module types.

An additional level of functional diversity arises from the intermodular "junction", i.e., the contact region between adjacent modules. This delimits the movement between neighboring modules and defines their relative spatial orientations. Thus, the shape, dimensions, and flexibility of a multimodular protein are dictated both by the nature of its component modules and the junctions between them. Moreover, a ligand-recognition site is commonly distributed over two, three, or four neighboring modules. As a consequence, the orientation of each module with respect to its neighbors may be critical for function. The nature of the junction between two modules is partly defined by the bulk and topology of the modules themselves and the length of linking peptide sequence (*4*). It is also a function of the representation of amino acids within the regions forming the intermodular interface and within the linker peptide. These residues are not always conserved—some examples of module pairs and higher multiples of modules for which structures are available to date reveal significantly different junction structures resulting in variations in intermodular angles and flexibility (*3*, *5*).

A case in point is the family of proteins concerned with the regulation of complement activation (RCA),[1] complement receptor type 1 (CR1), CR2, membrane cofactor protein (MCP), decay accelerating factor (DAF), C4 binding protein (C4bp), and factor H (*6*). All of these proteins are composed

† The Edinburgh Centre for Protein Technology is funded by the U.K.'s Biotechnology and Biological Sciences Research Council (BBSRC) and Department of Trade and Industry. M.D.K. is funded by the Human Frontiers Science Program. D.U. was funded by the Wellcome Trust. D.T.F.D. is funded by the Royal Society of Great Britain. The biomolecular calorimetry facility in Glasgow University is funded by the BBSRC. Work in St. Louis was supported in part by funding from the National Institute of Health (R01 AI41592) and from CytoMed Inc. (Cambridge, MA). J.P.A. and Washington University have a financial interest in CytoMed Inc.

* To whom correspondence should be addressed.
‡ The Edinburgh Centre for Protein Technology.
§ Washington University School of Medicine.
|| Institute of Cell and Molecular Biology.
⊥ Glasgow University.

10.1021/bi982453a CCC: $18.00
Published on Web 05/08/1999

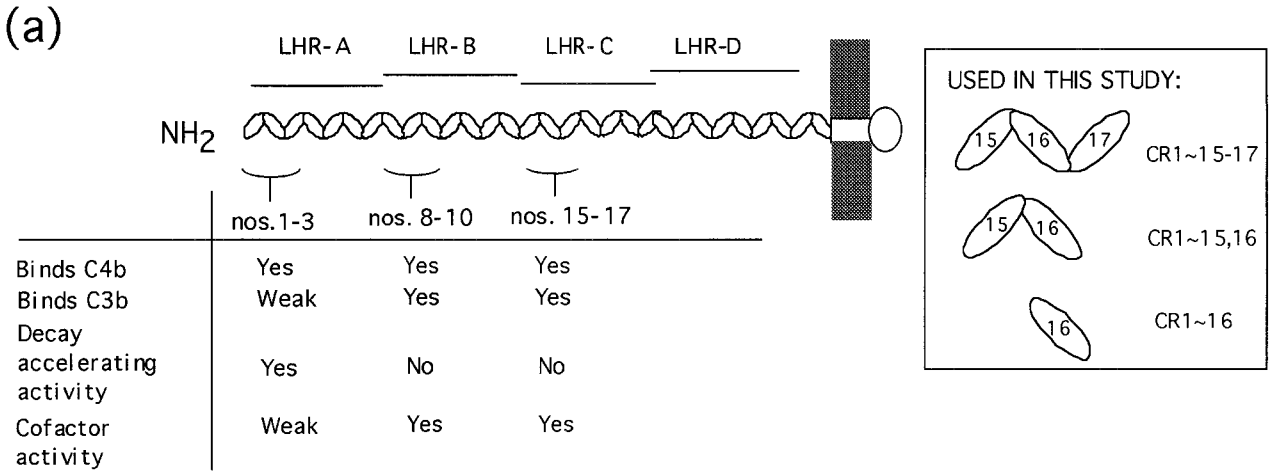
mainly from the "CP" (for complement protein—also called short consensus repeats or SCRs) module type. CP modules consist of approximately 60 residues and have four conserved cysteines disulfide bonded 1−3 and 2−4, a virtually invariant tryptophan, and several well-conserved glycines, prolines, and hydrophobic residues (*7*). Factor H has 20 CPs, CR1 has 30, CR2 has 15 or 16, C4bp b-chain has eight, and DAF and MCP have four each. A total of five structures of CPs have been solved following their recombinant expression as individual modules or as pairs of modules (*5*, *8*−*10*). In each case, the module has an elongated structure with the N-terminus and C-terminus lying at opposite poles of the long axis. A sheath of short β-strands encloses a hydrophobic core. The two disulfides lie at opposite ends of the core. The tryptophan side chain is mainly buried within the core and is situated adjacent to the disulfide which lies near the N-terminus. The two currently available examples of structures of CP module pairs display markedly different intermodular junctions, and it has proved impossible to model other CP junctions reliably. This is a barrier to interpreting the large body of mutagenesis data now available, particularly in the case of CR1 (*11*−*14*) and CR2 (*15*). A better understanding of structure−function relationships among the RCA proteins is an important step toward the design and assessment of more potent complement inhibitors with potential in a range of clinical settings (e.g., refs *16*−*18*) including xenotransplantation (e.g., ref *19*). Moreover, CR2, MCP, and DAF are all important viral targets (*20*−*22*). Thus, it is a priority to obtain information about the intermodular junctions and the angles between neighboring modules.

In this paper, we describe studies of a three-module fragment of CR1 (and some of its component modules) using solution techniques. The fragment under study here (see Figure 1a) comprises module numbers 15−17 (i.e., CR1∼15−17), a region which is almost identical in sequence and function to CR1∼8−10 (*23*). It is also similar to CR1∼1−3: the sequences of CR1∼15 and CR1∼1 are 55% identical; CR1∼16 and CR1∼2 are 70% identical; and CR1∼17 and CR1∼3 are 95% identical (see Figure 1b). CR1∼15−17 (and CR1∼8−10) bind to C3b and C4b and act as cofactors for factor I mediated proteolysis of C3b/C4b (*24*). CR1∼1−3 binds C4b (and C3b only very weakly) and accelerates the decay of the classical and alternative pathway convertase (*25*). The goal of the present studies was to establish the extent of intermodular contacts within CR1∼15−17 and compare our results with those of a similar study by Clark et al. on CR1∼1−3 (*26*). In contrast to their conclusion that CR1∼1−3 unfolds as a single unit, we have observed several discrete intermediates in the unfolding of CR1∼15−17 in equilibrium studies.

[1] Abbreviations: RCA, regulator of complement activation; CR1, complement receptor type 1; CP module, complement protein module; CR2, complement receptor 2; MCP, membrane cofactor protein; DAF, decay accelerating factor; C4bp, C4 binding protein; CR1∼*n*, CR1∼*n*, *m*, or CR1∼*n*−*m*, the *n*th module of CR1, or a fragment comprising the *n*th and *m*th modules of CR1, or a fragment comprising all modules from the nth to the mth modules of CR1; CR1∼*n*−*m*$^-$, CR1∼*n*−*m* with N-glycosylation sites mutated; CD, circular dichroism; DSC, differential scanning calorimetry; NMR, nuclear magnetic resonance; HSQC, heteronuclear single quantum coherence; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; GdnHCl, guanidine HCl; PCR, polymerase chain reaction.

EXPERIMENTAL PROCEDURES

1. Expression, Purification, and Assay of the CR1 fragments. For the production of the truncated, secreted forms of CR1, the *Pichia pastoris* system (Invitrogen, San Diego, CA) was used. The plasmid pSG5 (Stratagene, La Jolla, CA) containing cDNA encoding CR1∼15−21 [i.e., LHR−C (see Figure 1a), constructed as described previously (*14*)], served as a template for amplification of the DNA encoding CR1∼15−17 by polymerase chain reaction (PCR). The 5′ primer was GTGGCCTGGGGTCTCGAGAAAAGAGAGGCTGAAGCTCACTGCCAAGCAC and the 3′ primer was ATTTGGAGGCTCGAGTTAGTTAGGTAT. Each primer contains an *Xho*I site (underlined) for cloning into yeast expression vector, pPIC9 (*27*). In addition, the 3′ primer contains a translational stop codon. The amplified fragment was subcloned into the PCR II vector (Invitrogen). Following sequencing, which confirmed that the sequence encoding CR1∼15−17 from His901 to Asn1095 was correct, the cDNA was cloned into plasmid pPIC9. The *Xho*I site within the sequence for signal peptide of yeast α-factor was used for cloning. The yeast signal peptide at the amino terminus of the expressed CR1∼15−17 ensured that it was secreted into the growth medium. Transformation, selection of the transformants, and expression were performed according to the Invitrogen instruction manual. Briefly, plasmid linearized with *Bgl*II or *Sal*I was transformed into *P. pastoris* by electroporation using a Bio-Rad (Hercules, CA) Gene Pulser II under conditions recommended for yeast by Invitrogen. The yeast cells were plated on minimal medium without amino acids. This ensured selection of the transformants because untransformed yeast requires histidine. The transformants were tested for the production of CR1∼15−17 in 5 mL cultures. To assess levels of protein production, yeast was first grown in minimal medium containing 100 mM potassium phosphate, pH 6, 0.34% yeast nitrogen base with ammonium sulfate, without amino acids (Difco, Detroit, MI), 0.4% biotin, and 1% glycerol. After the yeast culture had reached an $OD_{414} > 6$, the cells were transferred to the medium containing 0.25% methanol instead of glycerol. The presence of methanol was necessary for protein expression because the CR1∼15−17 encoding region was under control of the alcohol oxidase promoter. The supernatants were screened for CR1∼15−17 by Western blotting using mAb 3D9 (*28*) which recognizes CR1∼17. Gels stained with Coomassie Blue showed that under reducing conditions the protein migrated as a very broad band of approximately 40 kDa instead of about 21 kDa as expected for a protein of approximately 200 amino acids. The slower than predicted migration indicated heavy glycosylation, and this was confirmed by treating a portion of the protein with endoglycosidase H. The broadness of the bands on SDS−PAGE suggested that the population of molecules was heterogeneous, probably due to nonuniform glycosylation. To remedy this problem, two N-glycosylation sites, one in CR1∼15 and the other in CR1∼16 (see Figure 1a), were removed by making mutations N918T and N987T using the QuikChange mutagenesis kit (Invitrogen). Threonine was chosen as a replacement for Asn because it is similar in size. The nonglycosylated protein (referred to hereafter as CR1∼15−17$^-$) migrated as a sharp band on SDS−PAGE at a position consistent with a molecular mass of about 20 kDa under nonreducing conditions and about 32 kDa under reducing



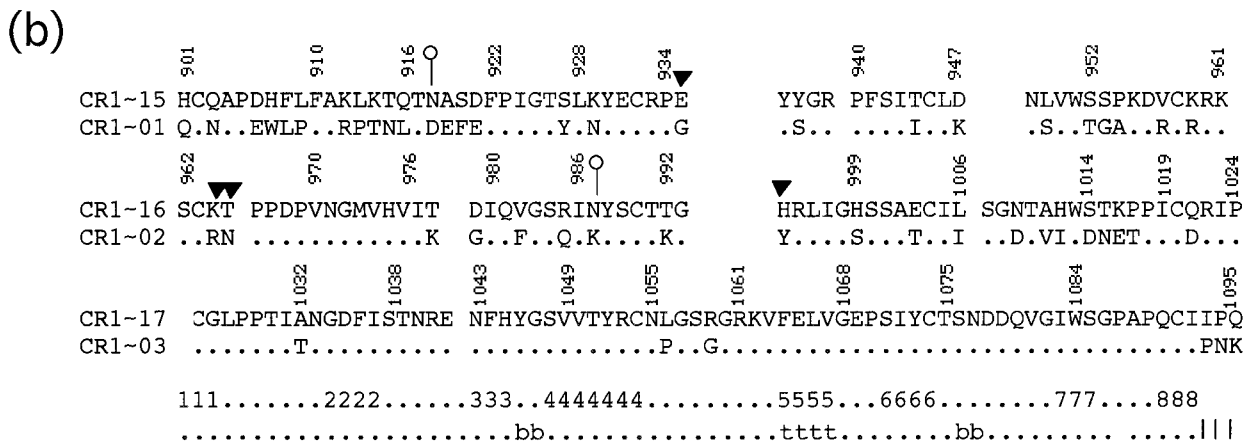


FIGURE 1: (a) A schematic representation illustrating (left-to-right) the N-terminus, 30 CP modules, transmembrane domain, intracellular domain and C-terminus of CR1. The four long homologous repeats (LHR∼A−D) into which the N-terminal 28 CP modules of CR1 may be divided are indicated above the schematic, and their biological functions are summarized below. Also shown (inset) are the names, abbreviations, and modular compositions of three protein fragments which have been overexpressed in *P. pastoris*, ^{15}N-labeled and used in this study. (b) Sequence alignment of each of module numbers 15−17 from CR1 with module numbers 1−3. Dots indicate identities. The white circles indicate Asn residues of CR1∼15−17 mutated to Thr in this study. The black triangles denote residues which are known to be critical for C4b binding by mutagenesis studies (see ref *14*). 111, 222, etc., indicate predicted β-strands numbered from 1 to 8. bb, tt, and ll indicate residues likely on the basis of homology to be involved in the intermodular junctions. b denotes the bottom of the module (i.e., the part that might be expected to contact the preceding module), t denotes the top of the module, and l denotes the short peptide linker.

conditions. The discrepancy between the estimated and the observed size under reducing conditions has been noted with other CR1 fragments, for example CR1∼8−14 (RH, MK, and JPA, unpublished observation).

The cDNAs encoding nonglycosylated CR1∼15,16 and CR1∼16 (i.e., CR1∼15,16$^-$ and CR1∼16$^-$) were constructed from the clone coding for CR1∼15−17$^-$. CR1∼15,-16$^-$ was derived from CR1∼15−17$^-$ by removing with *Eco*RI the sequences encoding for CR1∼17. CR1∼15,16$^-$ extends from His901 to Pro1024 and migrates at about 18 kDa under nonreducing conditions. CR1∼16$^-$ was amplified by PCR using CR1∼15−17$^-$ as a template. The 5′ PCR primer, TCTCTCGAGAAAAGAGAGGCTGAAGCTAAA-TCATGTAAAACTCCT contains an *Xho*I site. In the 3′ primer GCGGCCGCCCTAGGGTATTCGTTGACAAA-TTGC, a translational stop codon TAG is followed by a *Not*I site (underlined). The protein CR1∼16$^-$ starts with Lys961 and ends with Pro1024 and migrates at 8.5 kDa under nonreducing conditions.

The amino terminal sequence of each protein was confirmed by amino acid sequencing. In each case, the CR1 sequence was preceded by Glu-Ala-Glu-Ala, derived from the signal peptide of *P. pastoris* α-factor. The yield of the proteins was 30−100 mg/liter. On the basis of inspection of SDS−PAGE gels stained with Coomassie Blue, all of the expressed proteins were at least 80% pure. After expression, the cultures were centrifuged twice at 9000 rpm in a JA 10 rotor for 15 min at 4 °C to remove the yeast. The supernatants were filtered through a 0.22 μm Nalgene filter unit and concentrated 20−40-fold using Millipore (Bedford, MA) Centricon Plus 80 concentrators. The concentrated proteins were purified on a cation exchange column Bio-Rad UNO S6. To elute the proteins, a 0 to 1 M NaCl gradient in 25 mM Mes, pH 6.5, was used. CR1∼15−17$^-$ eluted at 0.23 M salt, CR1∼15,16$^-$ at 0.25 M salt, and CR1∼16$^-$ at 0.12 M salt. The eluates were concentrated to about 2 mL, the buffer was exchanged for 25 mM potassium phosphate, pH 6, and the samples were lyophilized.

For preparation of the ^{15}N-labeled sample, yeast was grown in the modified medium containing 0.34% yeast nitrogen base without amino acids and without ammonium sulfate,

supplemented with 1% ^{15}N-ammonium sulfate (Cambridge Isotopes Ltd., Andover, MA).

Secreted LHR C (consisting of modules 15−21 of CR1, see Figure 1) was expressed in COS 7 cells as described (*14*). Purified CR1∼15−17$^-$ and the COS cell supernatant containing LHR C were tested for binding to C4b and iC3 (C3 with a broken thioester bond which has activity similar to C3b) (*14*). Briefly, samples containing the proteins were incubated with iC3 or C4b conjugated to Sepharose in 25−100 mM NaCl. The bound proteins were eluted with 300 mM NaCl containing 1% Nonidet P-40 and quantified by sandwich ELISA using monoclonal antibody 3D9 as a capture antibody and a rabbit polyclonal to CR1 (29) for detection.

2. Physical Studies. Unless stated otherwise, all experiments were carried out in 25 mM sodium phosphate buffer, pH 6.4.

Differential scanning calorimetry (DSC) studies were conducted on a MC-2 differential scanning calorimeter (Microcal, Northampton, MA). The cell volume was 1.5 mL, rate of heating was 1 °C min^{-1}, and excess pressure was kept equal to 80 bar. The partial molar heat capacity and melting curve were analyzed using standard procedures (*30*). The data were processed using the software ORIGIN 2 (Microcal). The protein concentration was 0.5−4 mg mL^{-1} as determined by measurement of absorbance at 278 nm and according to calculated (*31*) extinction coefficients of 1.18 for CR1∼15−17$^-$, 1.07 for CR1∼15,16$^-$, and 0.94 for CR1∼16$^-$ for a 1 mg mL^{-1} solution in a 1 cm path length.

Measurements of circular dichroism (CD) were made using a Jasco-600 spectropolarimeter (Japan Spectroscopic Company, Tokyo) with cylindrical quartz cell of path length 0.02 cm at the EPSRC/BBSRC-funded center for CD at the University of Stirling with the kind assistance of Dr. Sharon Kelly and Professor Nick Price. The protein concentration was 0.2 mg mL^{-1}. All measurements were recorded at 25 °C except where stated. To obtain plots of percentage total change versus concentration of guanidinium HCl (GdnHCl) or temperature, the maximum ellipticity and readings at ±1.0 and ±2.0 nm of the maximum were averaged to smooth experimental errors in the curves. The concentration of stock GdnHCl solutions was confirmed using a refractometer and applying the formula $[M] = 57.2\Delta R + 38.7\Delta R^2 + 91.6\Delta R^3$, where R is the difference between refraction coefficient of buffer and that of the GdnHCl solution being assayed (*32*).

Fluorescence measurements were recorded on a Perkin-Elmer LS-50 spectrofluorimeter in a 100 μL cuvette of 0.3 cm path length at 20 °C unless otherwise stated. The excitation wavelength was 295 nm, and the emission spectra were recorded over 300−500 nm. The spectral bandwidth was 10 nm. The protein concentration was 0.3 mg mL^{-1}. The change in fluorescence intensity as a function of denaturant was analyzed by GRAFIT (Erithacus Software, Staines, U.K.) with equations (*33*) for one or two independent two-state unfolding transitions. The first transition for CR1∼15−17$^-$ was fitted to a single two-state transition while the second and third transitions were fitted to two consecutive two-state transitions.

Ultracentrifugation. Analytical centrifugation experiments were performed at the University of Nottingham (with the expert assistance of Professor Stephen Harding and Dr. Kornelia Jumel). Beckman (Palo Alto, CA) Optima XL-A and XL-I analytical centrifuges with absorption and Rayleigh interference optics were used. Sedimentation equilibrium experiments were run at 4 °C—double-sector cells with a 12 mm optical path length were loaded with 100 μL of 25 mM sodium phosphate, pH 6.5 buffer in the solvent channel and 80 μL of sample (0.5−4 mg/mL of CR1∼15−17$^-$, 25 mM phosphate buffer, pH 6.5) in the sample channel. For molecular mass analysis, MSTAR (absorption optics) and MSTARA (interference optics), which use the computerized M* method (*34*, *35*), were employed.

Sedimentation velocity experiments were performed using a 12 mm optical path length cell containing 400 μL of sample (as above) at a rotor speed of 55 000 rpm and at 4 °C. Sedimentation coefficients were evaluated from the absorption traces using the SVEDBERG algorithm (*36*). The average value for the Svedberg coefficient, $s^0_{20,w}$, was 1.89 ± 0.01 S, and for the frictional ratio, f/f_0, it was 1.48 ± 0.12. In addition, a range of hydration estimates was used in the Perrin function, P (or "frictional ratio owing to shape") (*37*). The average value of P was 1.33 ± 0.04. The axial ratio, a/b, was evaluated from P using the routine ELLIPS1 (*38*).

NMR measurements were performed on a Varian (Palo Alto, CA) INOVA NMR spectrometer operating at 600 MHz proton frequency, in a 5 mm triple-resonance pulsed field gradient probe. The ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) spectra of 0.3−0.6 mM samples of recombinant ^{15}N-labeled protein were acquired for 50 min each, at a range of temperatures from 37 to 65 °C, using a composite-pulse WATERGATE for water suppression (*39*). The acquisition times were 53 and 64 ms in t_1 and t_2, respectively. Eight scans were accumulated in each of 128 increments. The number of scans was increased to 196 per increment for experiments in the presence of 2.5−5 M GdnHCl, at 37 °C. The proton pulse width under these conditions was doubled in comparison with a 50 mM buffer, while the ^{15}N pulse width remained the same. Data were processed and plotted using FELIX software (MSI, San Diego, CA). The temperature of the sample in the probe was calibrated using ethylene glycol (*40*).

The cross-peaks in the HSQC spectrum of CR1∼15,16$^-$ at 37 °C were sequentially assigned on the basis of standard three-dimensional ^{15}N-edited nuclear overhauser and total correlation spectroscopy (see Supporting Information and Figure 2c). The HSQC spectra of CR1∼15,16$^-$ and CR1∼15−17$^-$ were compared to allow per-module assignments of cross-peaks in the spectra of the triple module. For example, if a pair of cross-peaks, one in the spectrum of CR1∼15,16$^-$ and the other in the spectrum of CR1∼15−17$^-$, had very similar chemical shifts ($\Delta\delta$N < 1 ppm; $\Delta\delta$H < 0.08 ppm), then it was assumed that the cross-peak in the spectrum of CR1∼15−17$^-$ arose from the amides of either the 15th or 16th modules. Cross-peaks in the CR1∼15−17$^-$ spectrum that did not correspond to cross-peaks in the spectrum of CR1∼15,16$^-$ were assumed to have originated from the amides of module 17. In the cases of series of experiments conducted at incrementally increasing temperatures or denaturant concentrations, the positions of cross-peaks were "followed" as they moved and per-module assignments made on that basis. It was only possible to do this for well-dispersed peaks. Although this approach is not as rigorous as a full sequential assignment under each set of conditions,

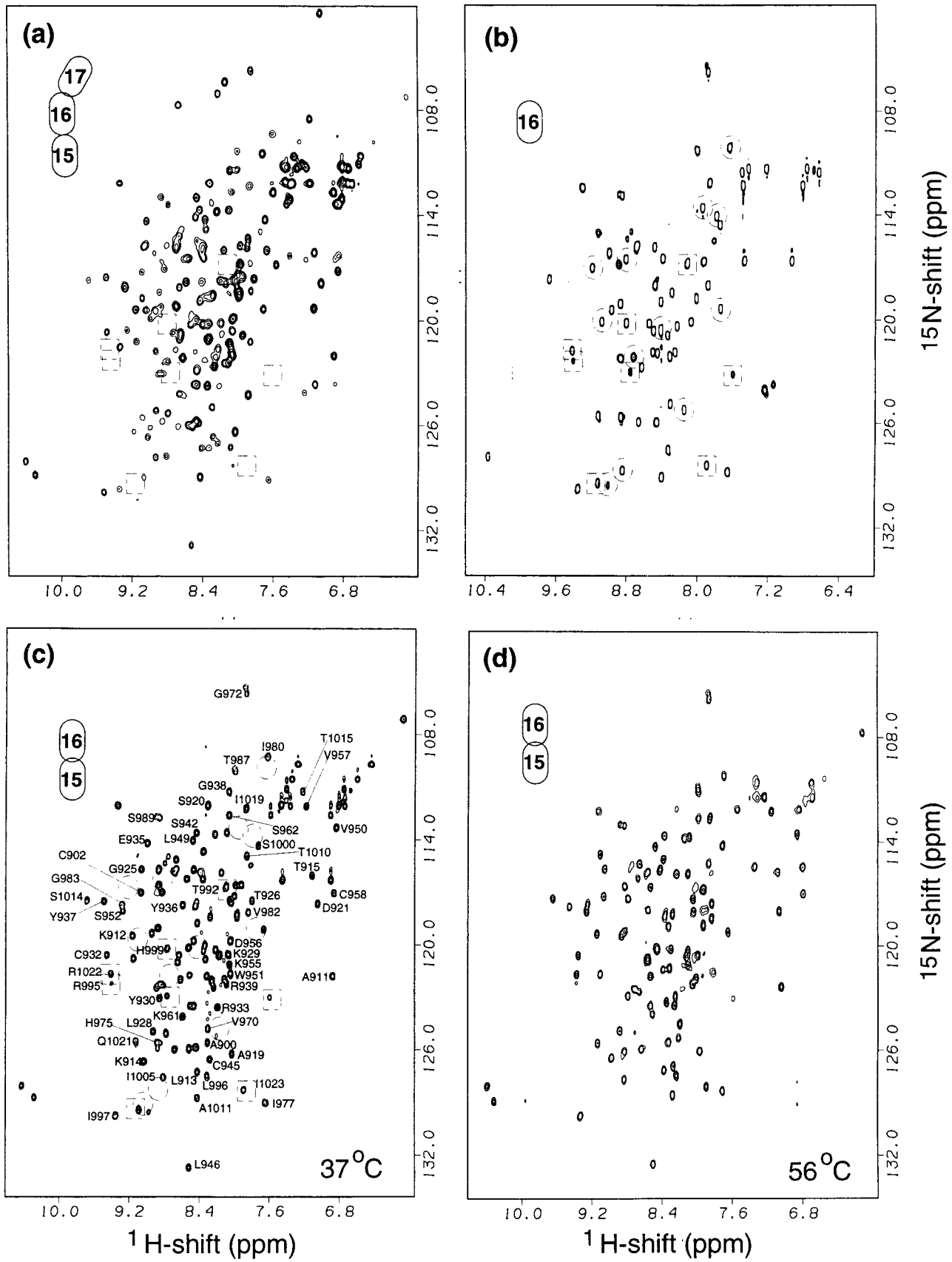


FIGURE 2: A series of ^{1}H,^{15}N-HSQC spectra collected on CR1∼16^{-}, CR1∼15,16^{-}, and CR1∼15−17^{-}. All spectra were collected at 37 °C and 50 mM phosphate buffer, pH 6.5 unless indicated. Protein concentrations were 0.3−0.6 mM. (a) The triple module fragment, CR1∼15−17^{-}; (b) The single module, CR1∼16^{-}; (c) The double module fragment, CR1∼15,16^{-}, the annotations indicate selected amino acid assignments. (d) CR1∼15,16^{-} recorded at 56 °C. The position of cross-peaks from CR1∼16^{-} that move by more than 0.08 ppm in the proton dimension, or 1.0 ppm in the ^{15}N dimension, or disappear when this module is joined to module 15 are indicated by circles in panels b and c; cross-peaks that move by the same margin or disappear when module 17 is also attached are indicated by squares in panels a and b.

Table 1: Binding Data

CR1 fragment	% binding[a] to iC3 at [NaCl]			% binding to C4b at [NaCl]		
	25 mM	50 mM	100 mM	25 mM	50 mM	100 mM
LHR C[b]	75 ± 12	62 ± 6	39 ± 11	83 ± 16	58 ± 9	7 ± 8
CR1∼15−17	78 ± 8	68 ± 10	41 ± 9	56 ± 12	42 ± 7	24 ± 10

[a] These data represent the mean ± SEM for three experiments. They refer to the percentage of protein binding to the two targets. [b] LHR C is composed of modules 15−21 from CR1 (see Figure 1)

it is extremely unlikely that more than a few individual assignments will prove erroneous and, for the purposes of this study, this level of reliability is sufficient. Cross-peaks of modules 15 or 16 which, upon addition of another module, were no longer within 0.08 ppm (^{1}H) or 1 ppm (^{15}N) of their original positions were used to infer information about module:module interactions.

RESULTS

1. Expression, Purification, Characterization and Biological Assay. The protein CR1∼15−17$^-$ binds iC3 and C3b (Table 1). Over a 4-fold range of ionic strengths, this ligand-binding activity is similar to that of the seven-module parental fragment, LHR C (i.e., CR1∼15−21). CR1∼15−17$^-$ is recognized by a monoclonal antibody to CR1, 3D9, and by a polyclonal antibody. Because all three CP modules, i.e., modules 15, 16 and 17, are required for activity, it was not possible to employ a functional assay to evaluate the structural integrity of CR1∼15−16$^-$ or CR1∼16$^-$. An ELISA assay could not be used because mAbs that recognize CR1 modules 15 or 16 are not available. However, CD, DSC, and NMR all confirmed the structural integrity of the nonglycosylated proteins (see below).

2. Hydrodynamics. Equilibrium sedimentation was used to determine the apparent molecular mass of CR1∼15−17$^-$ over a range of concentrations (0.5−4 mg/mL). The obtained value of 22.3 ± 1.0 kDa is close to that calculated from the amino acid sequence, 21.4 kDa. From values obtained for the Svedberg coefficient from sedimentation velocity measurements, for the Perrin function, and for the molecular mass, and assuming a typical level of hydration, an axial ratio of 6.0 ± 0.3 was calculated for CR1∼15−17$^-$.

3. Heteronuclear NMR Studies. A two-dimensional ^{1}H,^{15}N-HSQC spectrum correlates amide proton and nitrogen chemical shifts. Thus, in an HSQC spectrum of a protein, a cross-peak is observed for each nonproline 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 random coil polypeptide have a very narrow range of proton chemical shifts (*41*) and tend to overlap. Conformational exchange results in peak broadening or doubling. Thus, an HSQC spectrum provides a reliable method of assessing whether a protein is folded or not. Furthermore, the chemical shifts of cross-peaks arising from those amino acids that participate in interactions between domains or with other proteins provide a means of probing the nature and extent of domain−domain or protein−protein interactions.

A ^{15}N-labeled sample of CR1∼15−17$^-$ was prepared as described and examined using (^{1}H,^{15}N) NMR spectroscopy. Figure 2a displays the HSQC spectrum of CR1∼15−17$^-$, collected at 37 °C on a 0.4 mM sample, revealing approximately the expected number of well-dispersed, sharp cross-peaks consistent with fully folded, monodispersed protein (although at higher concentrations, a tendency toward aggregation was observed). Figure 2, panels b and c, display the HSQC spectra for the isolated 16th CP module of CR1 and for the double module fragment, CR1∼15,16$^-$, collected under similar conditions. Figure 2c indicates a representative selection of individual residue assignments of cross-peaks in the CR1∼15,16$^-$ spectrum achieved via a sequential backbone assignment (described in Supporting Information). Comparison of the HSQC spectra in Figures 2, panels a and c, enabled us to identify cross-peaks within the HSQC spectrum of the triple module fragment as originating from one or other of the component modules. The chemical shifts of many of the cross-peaks displayed by CR1∼16$^-$ alone appeared to be preserved in the spectra of CR1∼15,16$^-$ and CR1∼15−17$^-$. However, the positions of at least 20 cross-peaks, out of a total of about 60, have shifted (see Figure 2b). About 12 CR1∼16$^-$ cross-peaks appear to be perturbed by the attachment of module 15 (indicated by circles in Figure 2b), and eight cross-peaks are perturbed by the attachment of module 17 (indicated by squares in Figure 2b). There does not appear to be any displacement of the cross-peaks assigned to module 15 in CR1∼15,16$^-$ when module 17 is incorporated.

It was possible to assign approximately 30 cross-peaks to each of modules 15, 16 and 17, within CR1∼15−17$^-$ and subsequently to follow, in a qualitative fashion, the fate of individual modules within the triple fragment as the temperature was increased or the protein was denatured using GdnHCl. The results of such an experiment are shown in Figure 3 which is color-coded according to module. At 37 °C (refer back to Figure 2a), approximately 160 sharp cross-peaks may be distinguished, and there is no evidence of any unfolded protein in the sample. At 42 °C (data not shown), all modules remained intact, but a very small peak appeared at 10 ppm (^{1}H shift). This may be reasonably assigned to the ring NH of Trp in random coil conformation and is evidence for a very minor degree of module unfolding. As the temperature was increased to 47 °C (data not shown), the amount of unfolded material remained very small. At 52 °C (Figure 3a), the intensity of the random coil Trp peak (indicated by arrow) had increased, and a number of broad overlapping resonances appeared toward the center of the spectrum (green in the figure). However, a substantial proportion of all three modules remained largely intact as judged by the persistence of strong cross-peaks from each of the three modules with only a few changes in chemical shift. It was not feasible to quantify percentages of folded and unfolded material due to complications of line width and chemical or conformational exchange rates. However, by inspection of representative peak intensities, it appears that module 17 is more unfolded than modules 15 or 16. At 57 °C (Figure 3b), the intensity and number of the random coil cross-peaks increased greatly. It is evident that at this temperature module 17 had become unfolded in a great majority of the molecules, as judged by the absence of its cross-peaks. Modules 15 and 16 were still intact within a substantial fraction of the population, as indicated by the

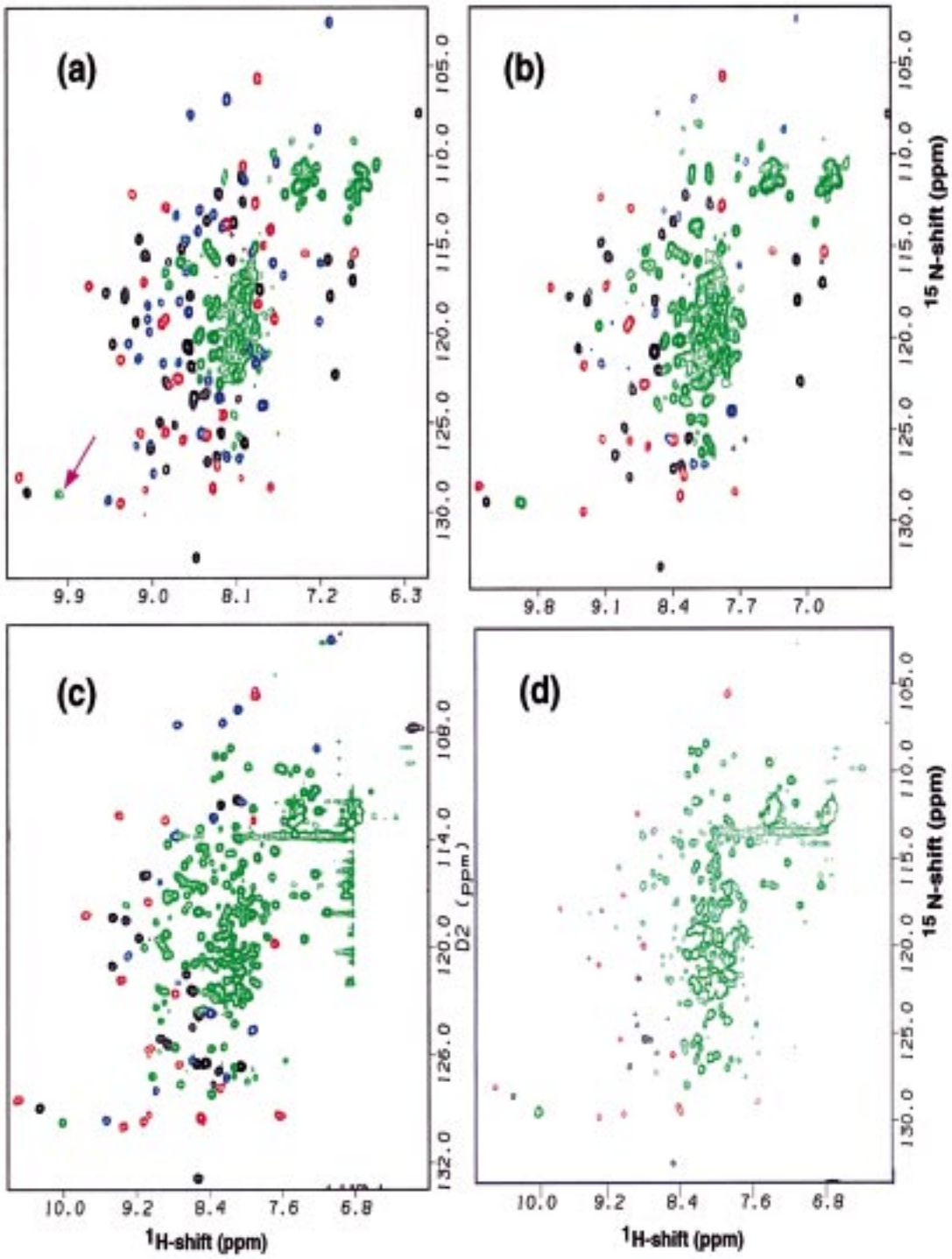


FIGURE 3: A series of ^{1}H,^{15}N-HSQC spectra collected on 0.3 mM CR1∼15−17$^-$ color coded according to module. Cross-peaks assigned to module 16 are in red, to module 15 in black, and to module 17 in blue. Cross-peaks where per module assignment is difficult because they are in a crowded position or because they correspond to random-coil positions, are in green. Spectra were recorded at (a) 52 °C, (b) 57 °C, (c) 37 °C but in the presence of 2.5 M GdnHCl, (d) 37 °C but in the presence of 4 M GdnHCl.

presence of about 20 dispersed cross-peaks for each of these two modules—although the cross-peaks are less intense. At 65 °C (data not shown), extensive precipitation occurred and a 1D spectrum collected rapidly during the precipitation event indicated an absence of folded structure. However, upon returning the sample to 42 °C, much of the protein redissolved and, according to a further HSQC experiment (data not shown), had refolded. The thermal stability of modules 15 and 16 was also monitored by recording an HSQC spectrum of the double module CR1∼15,16$^-$ at 56 °C. As shown by comparison of Figure 2, panels c and d, there is little or no evidence of module unfolding upon the temperature being raised from 37 to 56 °C.

The unfolding of CR1∼15−17$^-$ induced by the presence of GdnHCl was also monitored using the HSQC technique. On the basis of fluorescence data (see below), 2.5, 3, 4, and 5 M concentrations of GdnHCl were selected as representing informative stages during the course of denaturation. At 2.5 M GdnHCl (Figure 3c), a random-coil Trp peak is evident. On the basis of the number of identifiable cross-peaks (19

for module 15, 17 for module 16, and only 14 for module 17) and their relative intensities, a larger proportion of module 17 had unfolded than was the case for modules 15 and 16. This trend continued at 3 M GdnHCl (data not shown) where the random-coil Trp peak had more than doubled in intensity, cross-peaks from module 17 were weak, and cross-peaks from modules 15 and 16 became broader. As is apparent from Figure 3d, module 17 was unfolded at 4 M GdnHCl and there was considerable unfolded material present in the sample. However, a small proportion of modules 15 and 16 remained intact. At 5 M denaturant (data not shown), all the chemical shift dispersion was lost, indicating that, for the entire population of molecules in the sample, all three modules had lost their tertiary structure. This process was reversible as indicated by an HSQC spectrum (data not shown) collected on a sample that had been exposed to 7 M GdnHCl and then dialyzed to remove the denaturant.

4. Differential Scanning Calorimetry Studies. To further investigate the thermal unfolding of CR1∼15−17$^-$, a study was undertaken using DSC (Figure 4). In a plot of specific heat capacity versus temperature, the triple module exhibits an asymmetric profile which may reasonably be fitted to three overlapping melting transitions, as shown (Figure 4a), with midpoints at 50.5, 57.5, and 61.5 °C. In all cases there is imperfect agreement between calorimetric and van't Hoff enthalpies.

To investigate further, DSC profiles were collected on the single module CR1∼16$^-$ and on the module pair CR1∼15,16$^-$. The plot for CR1∼16$^-$ shows a single broad transition with a midpoint of 69 °C (Figure 4c). As with the triple module, there is an imperfect match between calorimetric and van't Hoff enthalpies, indicating that the transition is not two-state. CR1∼15,16$^-$ yielded a more complex profile in which the main peak has a distinct shoulder (Figure 4b). The melting profile may be fitted to two transitions, at 54 and 65 °C. Once again, neither transition displays well-matched calorimetric and van't Hoff enthalpies. The calorimetric enthalpies and melting temperatures are summarized in Figure 7.

5. Circular Dichroism Studies. Far-UV CD spectra of proteins are dominated by peptide bond dichroism and are therefore sensitive to secondary structure, while near-UV CD spectra reflect the environments of aromatic side chains and contain contributions from the disulfide bonds (*42*). By measuring CD spectra as a function of temperature and GdnHCl concentration, it is possible to complement the NMR, DSC, and fluorimetry studies. As has been observed previously (*43*) CD spectra of CP modules are characterized by a positive ellipticity in the 220−240 nm region. This could arise from the presence of tryptophans as has been noted for human carbonic anhydrase by Fresgard et al. (*44*), who showed that Trp residues in β-strands contribute positively in this region of the spectrum.

Figure 5, panels a and b, show CD spectra obtained for CR1∼15−17$^-$ obtained at various temperatures. There was little loss of structure below 40 °C, while the major transition lays between 50 and 60 °C. The outcome of the variable temperature far-UV CD studies on CR1∼15−17$^-$ is compared to the results of similar experiments on double- and single-module fragments in Figure 5c. However, aggregation and precipitation competed with unfolding at temperature above 65−70 °C, rendering interpretation difficult. A more useful study was carried out using increasing concentrations of GdnHCl as a denaturant and recording far-UV spectra (Figure 5d). For all of the fragments, there was little loss of secondary structure below 2.5 M GdnHCl and total loss of secondary structure at 6 M GdnHCl.

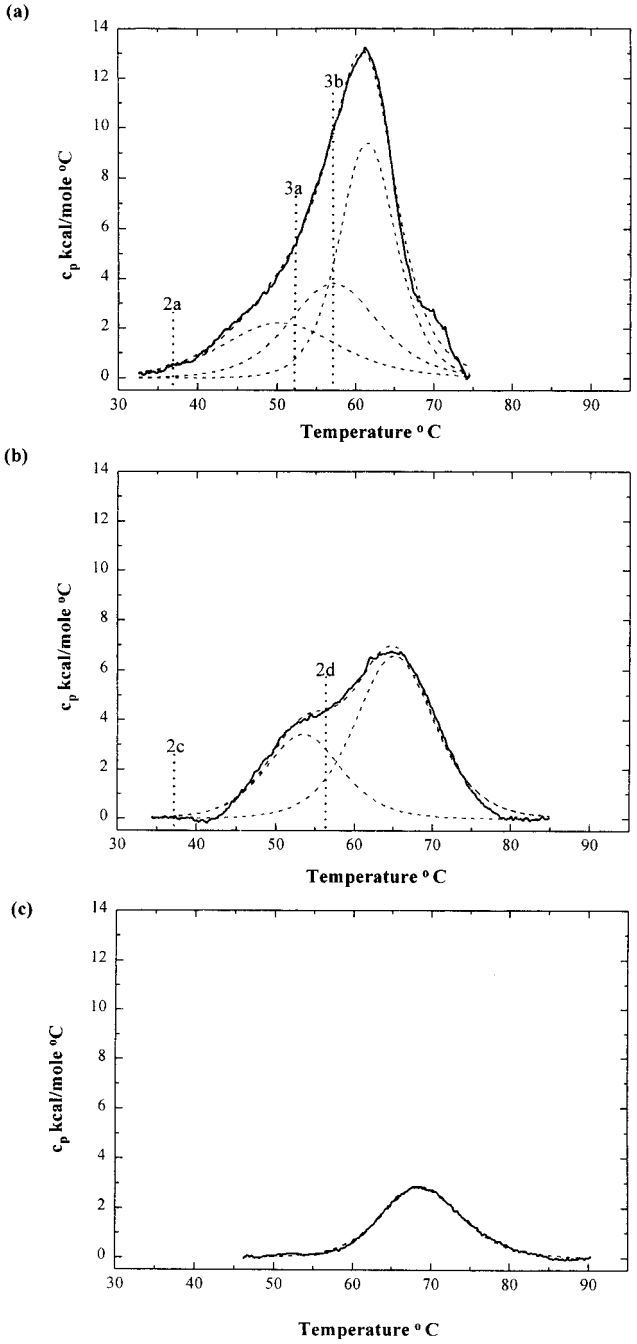


FIGURE 4: Differential scanning calorimetry studies on (a) CR1∼15−17$^-$ (b) CR1∼15,16$^-$, and (c) CR1∼16$^-$. Experimental data after baseline subtraction are shown by the solid lines. Dashed lines indicate the computer-fitted data. The annotations, (2b), (2d), (3a), etc., cross-refer to Figures 2 and 3 and correspond to the temperature at which the appropriate HSQC spectra were recorded.

6. Fluorescence Studies. Each module has a conserved Trp, and there are no Trp residues elsewhere in the sequence (Figure 1b). On the basis of homology with experimentally derived structures, the conserved Trp side chains can be expected to occupy largely buried positions within the hydrophobic cores of the modules, adjacent to a disulfide

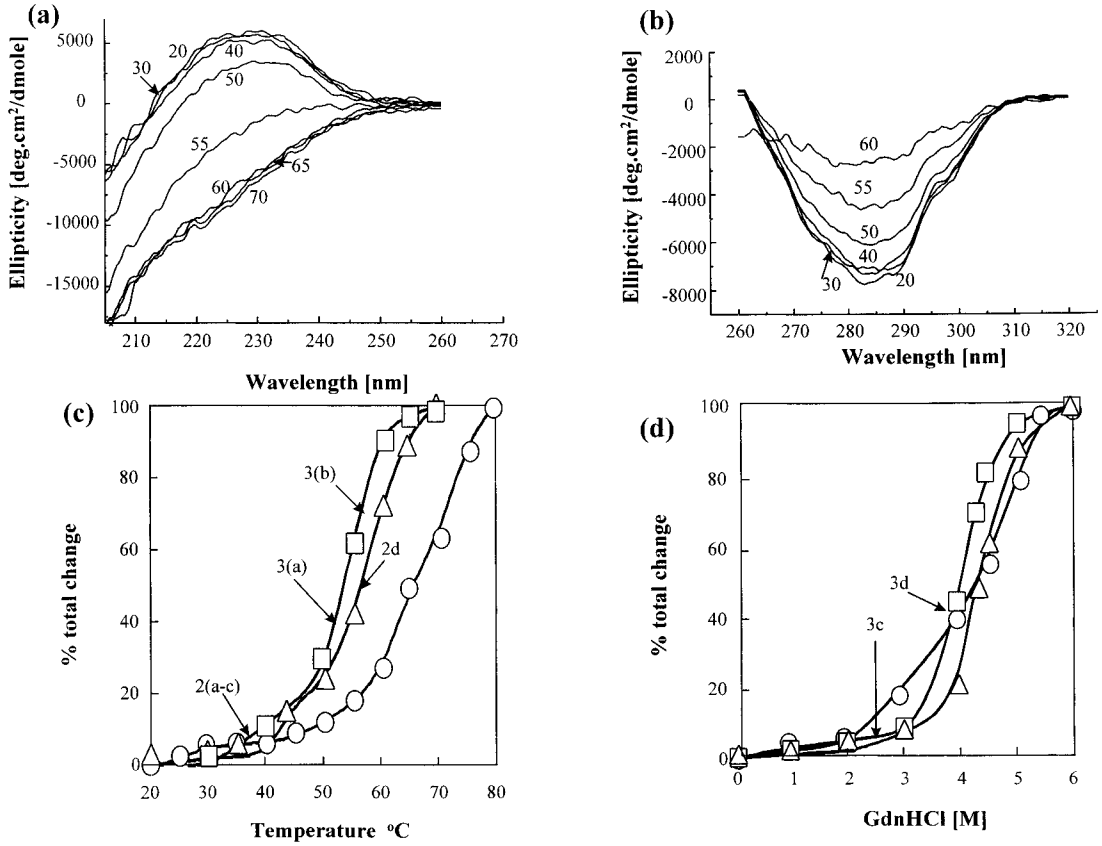


FIGURE 5: Circular dichroism spectra. A series of (a) far-UV and (b) near-UV CD spectra for CR1∼15−17⁻ were collected over a range of temperatures, as indicated by the numbers. (c) Thermally induced denaturation, as monitored by far-UV CD, of (□) CR1∼15−17⁻; (△) CR1∼15,16⁻; (○), CR1∼16⁻. This plot was extracted from the raw CD data as described in the text. (d) GdnHCl-induced denaturation as monitored by far-UV CD; symbols as above. The annotations, (2a), (2b), (3a), etc., cross refer to Figures 2 and 3 as before.

and close to the N-terminus (i.e., the junction with the previous module).

Excitation at 295 nm resulted in a fluorescence emission spectrum with a maximum at 340 nm. Upon denaturation of CR1∼15−17⁻ in 7 M GdnHCl there was an increase in fluorescence emission intensity at 340 nm of about 14-fold, a consequence of solvent exposure of the Trp side chain, and a shift in the maximum to 351 nm. The increase in the intensity of tryptophan fluorescence upon denaturation was smaller for CR1∼15,16⁻—about 10-fold—and smaller still for the single module CR1∼16⁻—about 7-fold. Fluorescence emission intensity at 340 nm upon excitation at 295 nm was subsequently used to monitor unfolding of modules within CR1∼15−17⁻. The fluorescent intensity changes were reversible.

Below 50 °C, exposure of Trp side chains to solvent was minimal for single, double and triple module fragments (Figure 6a). A single transition was observed in each case in the 55−70 °C range. However, Trp fluorescence is known to be quenched by heating, and the protein precipitated at 70−75 °C. A more useful experiment employed GdnHCl as a denaturant (Figure 6b). These data are represented as a semilog plot to emphasize the changes at low intensity. Using this representation, it is clear that in the case of the triple domain there are three clearly resolved transitions. Figure 6b indicates that there are two resolved transitions for CR1∼15,16⁻, which correspond closely to the second and third transitions for CR1∼15−17⁻. CR1∼16⁻ shows a broad melting profile which can be fitted to two overlapping transitions. Conformational stabilities and midpoints of unfolding were extracted as described and are summarized in Figure 7.

DISCUSSION

The recombinant protein, CR1∼15−17⁻, is a biologically active protein fragment containing three CP modules joined by two short peptide sequences. Our aim was to probe the extent and stability of intermodular junctions using a range of physical techniques but without recourse to a full structure determination. The data obtained are consistent with a model of CR1∼15−17 in which the modules are joined end-to-end with little or no direct interaction between modules 15 and 17. The interface between modules 15 and 16 is extensive and stable while that between modules 16 and 17 is smaller. The 16−17 junction unfolds more readily than the 15−16 junction, and module 17 unravels independently of modules 15 and 16 which are relatively stable. The evidence for such a model (summarized schematically in Figure 7) is set out below.

The Modules Are Joined in an End-to-End Fashion. Most of the ^{15}N and ^{1}H chemical shifts of module 16 are the same whether the module is expressed singly or as the central component of a triple module fragment. Since such chemical shifts are sensitive to changes in structure, the spectra in Figure 2 demonstrate that the overall structure of module 16 remains largely unaltered upon coupling to modules at both its N- and its C-terminus. A similar result was obtained for modules 15 and 16 of factor H (*10*). Furthermore, the chemical shifts of CR1∼15 are not perturbed by the presence

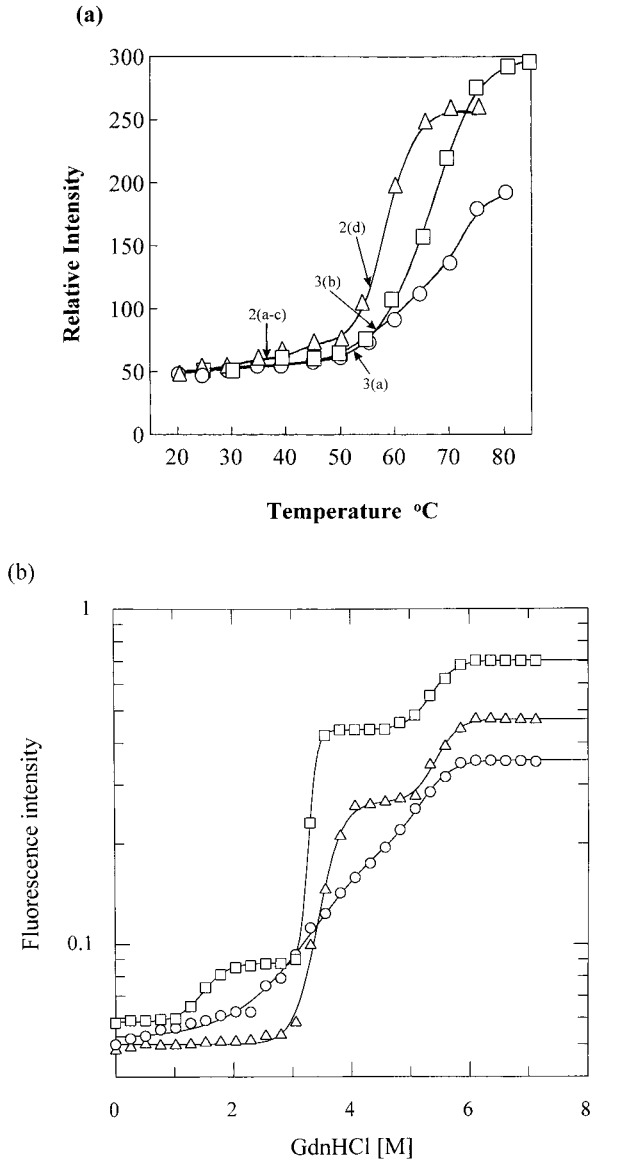


FIGURE 6: Fluorescence studies on CR1∼16⁻, CR1∼15,16⁻, and CR1∼15−17⁻. (a) Fluorescence intensity at 340 nm as a function of temperature for (□) CR1∼15−17⁻; (△) CR1∼15,16⁻; (○) CR1∼16⁻. The curves shown are for clarity only. The annotations (2a), (2b), (3a), etc., cross-refer to Figures 2 and 3 as before. The curves were fitted to the data as described. (b) Fluorescence intensity vs [GdnHCl] for (□) CR1∼15−17⁻; (△) CR1∼15,16⁻; (○) CR1∼16⁻.

of CR1∼17, suggesting a minimum of interaction between these two modules. These results are consistent with the modules of CR1∼15−17⁻ being joined end-to-end, as seen in previous examples (*5*, *10*). This is not to say that the modules are not tilted with respect to one another, rather the overall path taken by the three modules is roughly linear. In agreement with this conclusion, our ultacentrifugation study indicated that CR1−15−17⁻ had a highly elongated conformation with an axial ratio of between 5.7 and 6.3. That strings of CP modules form elongated structures is also indicated by investigation of factor H (20 modules) and CR2 (15 modules) using electron microscopy (*45*, *46*) and of factor H by neutron scattering (*47*).

The 15−16 Junction Is More Extensive Than the 16−17 Junction. While most amide chemical shifts of CR1∼16⁻ are not perturbed by neighboring modules, about one-third of them show some changes. This is consistent with the idea that neighboring modules approach one another closely forming an intermodular junction with accompanying changes in chemical shifts of residues in the junction. That 12 amides in CR1∼16⁻ are affected when module 15 is attached, compared to eight when module 17 is added, is consistent with more extensive contacts between modules 15 and 16 than between modules 16 and 17, under these conditions.

The 15−16 Junction Has a Higher Conformational Stability Than the 16−17 Junction. According to our model, the most unstable component of CR1∼15−17⁻ is the 16−17 junction. Evidence for this derives from both GdnHCl and temperature studies. Fluorescence experiments show a small transition with a midpoint at 1.5 M GdnHCl, absent in CR1∼16⁻ and CR1∼15,16⁻. There is only a small increase in fluorescence so this cannot represent the unfolding of a module. The HSQC spectrum at 2.5 M GdnHCl confirms that all three modules are largely intact, although module 17 has begun to unfold. Far-UV CD demonstrates very little loss of structure below 2 M GdnHCl. The 1.4 M transition must therefore correspond to the melting of an intermodular junction. The DSC profile shows a transition at 50.5 °C, which must also correspond to melting of a junction since there is very little module unfolding below this temperature according to fluorescence, NMR, and CD. That the 1.4 M GdnHCl and 50.5 °C transitions correspond to the melting of the 16−17 junction is confirmed by the observation that both transitions are missing from the respective DSC and fluorescence plots for CR1∼15,16⁻ and CR1∼16⁻.

The next most unstable entity is the 15−16 junction. Evidence for this comes from comparison of HSQC spectra of CR1∼15,16⁻ at 37 and 56 °C. There is very little change in amides chemical shifts over this range, implying very little module unfolding. Measurements of fluorescence also indicate little decrease of ordered structure in the double module fragment below 55 °C. However, CD does indicate loss of structure in this temperature range, and the DSC profile for CR1∼15,16⁻ shows an obvious transition at 54 ± 0.5 °C. Since this cannot correspond to the melting of a module, it must arise from denaturation of the 15−16 junction. The 54 °C transition presumably corresponds in part to the transition observed by fluorescence measurements for CR1∼15,16⁻ centered at 3.6 M. However, the 3.6 M transition is associated with a substantial change in fluorescence intensity and loss of structure according to both NMR and CD. Furthermore, an equivalent transition (3.6 M GdnHCl) is observed in fluorimetry of CR1∼16⁻. Therefore, the 3.6 M transition cannot correspond purely to junction melting and must also encompass module melting. A 57.5 °C transition is also present in the DSC profile of CR1∼15−17⁻, and a 3.3 M GdnHCl transition appears in the fluorescence measurements for the triple module.

A problem with this analysis is the value of 181 kJ mol^{-1} for ΔH estimated for the 50.5 °C transition in the DSC profile of CR1∼15−17⁻ (see Figure 7). This seems high when compared with the value for the 54 °C transition of CR1∼15,16⁻ (185 kJ mol^{-1}) corresponding to the unfolding of the more stable 15-16 junction. It also seems high in comparison to the 57.5 °C transition of CR1∼15−17⁻ (239 kJ mol^{-1}) corresponding to melting of both the 15−16 junction and module 17 (see below and Figure 7). An explanation could lie in the difficulty in deconvoluting the

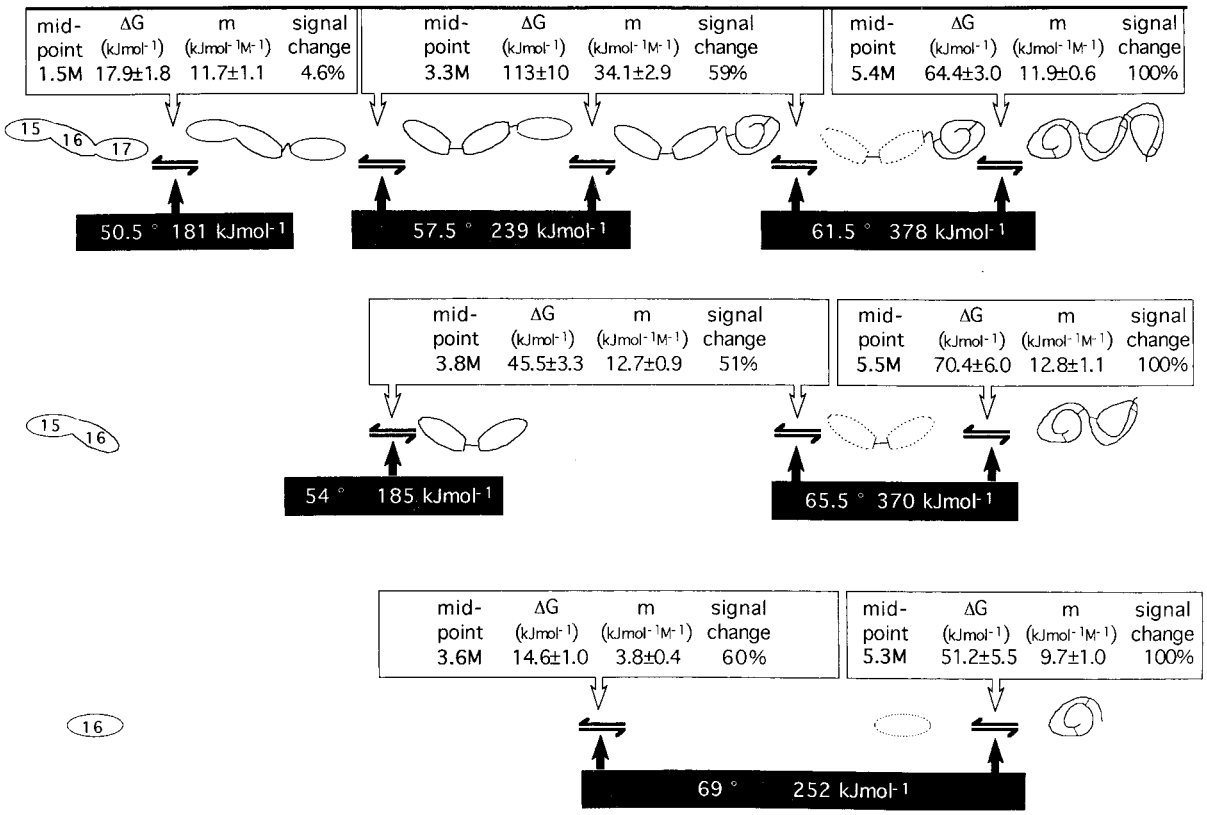


FIGURE 7: Combined schematic and data table for melting of CR1 fragments. A schematic is used to illustrate the proposed intermediates obtained in these studies. Transitions are represented by double arrows. The modules represented with dotted lines are probably in an intermediate folded state. The white boxes show data derived from plots of fluorescence intensity vs [GdnHCl] (see Figure 6b) as described. The midpoint, free energy change associated with unfolding in water, value of *m* (slope at transition) and the cumulative percentage of total change in fluorescence intensity (i.e., at 7 M GdnHCl) is given for the transition(s) indicated directly beneath each box. The black boxes contain melting points and calorimetric enthalpies for the transitions indicated above each box. Some of the data boxes cover two or three putative transitions which cannot be resolved by that particular experimental technique.

multiple overlapping transitions within the CR1∼15−17$^-$ DSC profile. For example, it seems unlikely that the 15−16 junction is affected by the presence of module 17, yet no 54 °C transition is apparent for the triple module. It is possible that a fourth transition at this temperature is present within the DSC profile but not revealed by the fitting algorithm, rendering unreliable the calculated values of ΔH and melting point for the first two fitted transitions.

Module 17 Is the Least Conformationally Stable of the Three Modules. The HSQC experiment (see Figure 2a) confirms that each of the modules in CR1∼15−17$^-$ forms a stable entity at 37 °C in the absence of denaturant. At 52 °C there is evidence of unfolded material. At 57 °C, no cross-peaks from folded module 17 are detectable. The HSQC spectra of CR1∼15,16$^-$ (Figure 2, panels c and d) imply practically no unfolding of these modules below 56 °C. Therefore, module 17 is partially unfolded at 52 °C and completely unfolded at 57 °C. The DSC profile of CR1∼15−17$^-$ is consistent with a transition at 57.5 °C, which could correspond to melting of module 17; although (see above), it might also include melting of the 15−16 junction. According to CD, there is a major loss of structure between 50 and 55 °C for CR1∼15−17$^-$ which could correspond to module 17, but it is difficult to compare meaningfully percentage loss of structure in the thermal CD studies due to precipitation and aggregation at higher temperatures.

In experiments using CD to monitor denaturation by GdnHCl, there is a much bigger loss of structure in CR1∼15−17$^-$ than in CR1∼15,16$^-$ over the 3−4 M range—again consistent with the melting of module 17. The NMR data support this: module 17 is completely unfolded at 4 M GdnHCl while modules 15 and 16 still yield some dispersed cross-peaks. The fluorescence data for CR1∼15−17$^-$ show a strong increase in intensity with a midpoint of 3.3 M GdnHCl. This transition provides a major contribution to the overall enhancement in intensity upon denaturation of CR1∼15−17$^-$ and is much larger than the equivalent transition in the module pair, CR1∼15,16$^-$. The difference must arise purely from melting of module 17 since the 16−17 junction had denatured at 1.4 M GdnHCl. It is consistent with a strongly quenched Trp in module 17 becoming completely exposed to solvent. It is interesting that the overall enhancement (per milligram of protein) in Trp fluorescence upon denaturation is much stronger for the triple module than for the double or single modules. It implies that the fluorescence emission of the Trp side chain of each of the 3-folded modules is not all quenched to the same extent. That the Trp side chains are in distinct environments is borne out by the HSQC spectra which show that the Trp of modules 15 and 16 are shifted downfield of the random coil position while the Trp of module 17 is shifted upfield.

Modules 15 and 16 Have About Equal Conformational Stability. The DSC profile for CR1∼15,16$^-$ shows two discrete transitions as does the GdnHCl/fluorescence plot. In the absence of other data, it would be reasonable to assign one transition to each of the two modules. However, the

NMR data are unambiguous—there is no evidence that modules 15 and 16 unfold separately and the first DSC transition does not correspond to the loss of a module. The second DSC transition (65.5 °C) corresponds to melting of both modules 15 and 16 (that two components are melting might explain why the calorimetric value of ΔH is bigger than the van't Hoff values of ΔH). In CR1∼15−17^{-}, the equivalent transition is at 61.5 °C (see Figure 7), implying the modules are destabilized by addition of module 17. Module 16 alone has a melting point at 69 °C, which implies that attachment of module 15 destabilizes module 16. The calorimetric transitions become progressively narrower going from single to double to triple module fragments. This implies an increase in the cooperativity of unfolding as the number of modules increases. The explanation for this presumably derives from junctions being less stable than modules. Consequently, the junction between modules 15 and 16 will denature prior to the modules themselves. Under these circumstances, module 16 might become destabilized due to the presence of the other module which is now unconstrained, mobile, and exerting a strain upon the linker region due to its inertia. The residue immediately prior to the first Cys residue of CP modules normally participates in β-sheet formation with residues within the modules, and this might be disrupted. Interestingly, our observations contrast with those for CR1∼1−3 where module 3 was seen to be stabilized by its junction with module 2. The explanation could lie in CR1∼1−3 having tighter junctions and less stable modules.

An HSQC spectrum collected on CR1∼15−17^{-} at 4 M GdnHCl, which lies between the two transitions, reveals that equal proportions of both modules 15 and 16 are unfolded. At 5 M GdnHCl, at the start of the second transition, no properly folded modules are detectable in the HSQC spectrum. A lack of secondary structure at 5 M GdnHCl is likewise indicated by CD. This implies that the two modules are in a molten form and sufficiently mobile to destroy the chemical shift dispersion and ellipticity, although tryptophans are still partially buried in a hydrophobic cluster. Studies of module 16 alone support the idea of an intermediate; the fluorescence studies show a broad profile for CR1∼16^{-} denaturation which may be fitted to two overlapping transitions. These correspond reasonably well to the better resolved transitions seen in CR1∼15,16^{-}. The nature of this intermediate requires more investigation but it may be dependent upon the disulfides linkages.

Free Energies of Unfolding May Be Extracted from the Fluorescence Data. Since some discrete reversible transitions are observed in the denaturation of CR1∼15−17^{-} by GdnHCl, it is feasible to extract values for ΔG between folded and unfolded conformations of the components of CR1∼15−17^{-} (see Figure 7). The overall free-energy changes, as determined by fluorescence spectroscopy, for complete denaturation are 66 ± 7, 116 ± 9, and 195 ± 15 kJ mol^{-1} for CR1∼16^{-}, CR1∼15,16^{-}, and CR1∼15−17^{-}, respectively. Each module in CR1 contains approximately 66 amino acids and two disulfide bonds. The stability of module 16 is comparable to that of bovine pancreatic trypsin inhibitor, a small protein containing 56 amino acids, and three disulfide bonds (*39*). These measurements of stability contrast with the work of Clark et al. (*26*) who found that ΔG of a single CP module was only about 35% smaller than the ΔG of a triple CP module fragment (i.e., CR1∼1−3). Furthermore, the stability of the proteins in the Clark et al. study was much lower, similar in magnitude to the free energy change for the first transition in CR1∼16^{-} and CR1∼15−17^{-}.

Clark et al. had access to module 3 in isolation, which has an almost identical sequence to module 17 (Figure 1b). The enhancement in fluorescence upon unfolding of CR1∼3 was 14-fold and the same enhancement was observed for CR1∼1−3, implying that, in all three modules, the Trp was quenched to the same extent. The third module in isolation (midpoint, 2.5 M GdnHCl, ΔG of unfolding, 13 kJ mol^{-1}) is less stable than module 17 in the context of CR1∼15−17^{-} (midpoint 3.3 M GdnHCl, apparent $\Delta G \approx 60-65$ kJ). However, CR1∼1−3 appears to melt as a single unit with a midpoint of 4.8 M GdnHCl and ΔG of ∼20 kJ mol^{-1} and showed no evidence for discrete transitions that could be ascribed to particular modules or junctions. One explanation for this is that both the 1−2 and the 2−3 junctions are well structured while the modules in isolation are not very stable yielding a highly cooperative system. In contrast, CR1∼15−17^{-} appears to have junctions with varying conformational stabilities. Selective melting of CP modules was also observed in factor H (*10*), in the vaccinia virus complement control protein (PNB, unpublished observation), and in an earlier calorimetric study of C1r (*48*).

Functional consequences. In light of the different models proposed for CR1∼15−17^{-} and CR1∼1−3, it is interesting to note the differences in sequence between these two fragments. On the basis of sequence alignment, residues likely to be involved in formation of the junction are summarized in Figure 1 where b indicates the bottom face of the module (i.e., the junction with the previous module), t indicates the top face, and l indicates the short peptide linker. The bottom faces of the 17th and third modules are probably identical but the top faces of the 16th and second modules differ in that His999 in module 16 is replaced by the more typical Tyr in module 2. Such a change could effect the structure of the junction. The structural consequences of this change have not been determined experimentally, but the functional consequences have been examined as part of a broader mutagenesis study. Krych et al. (*11*, *12*) found that swapping this Tyr for a His significantly decreased the ability of CR1∼1−3 to bind C4b. On the other hand replacing the His in CR1∼9 (identical to CR1∼16) with Tyr enhanced the ability of CR1∼8−10 (identical to CR1∼15−17) to bind to C4b.

The studies reported here were conducted on a nonglycosylated fragment of CR1. Although this fragment is biologically active, the influence of glycosylation, and the attachment of modules 14 and 18 on its melting properties, will be the subject of future investigations.

In summary, we have expressed nonglycosylated single, double, and triple module fragments of CR1 and shown they are joined in the expected end-to-end fashion. A combination of physical techniques have allowed us to dissect the denaturation profiles of these fragments. The melting behavior of the biologically active triple module fragment is a complex function of the intrinsic stability of the component modules and the extent to which the intermodular junctions are structured. A model is presented (Figure 7). These studies have demonstrated some interesting differences between two

binding sites (CR1∼1−3 and CR1∼15−17) which are similar in sequence and have distinct but related biological roles.

ACKNOWLEDGMENT

We are grateful to Professor Nick Price and Dr. Sharon Kelly of the BBSRC-funded circular dichroism facility at Stirling University for their invaluable expert assistance with the CD studies. We are also grateful to Professor Stephen Harding and Dr. Kornelia Jumel for their essential help with recording and analyzing the ultracentrifugation data.

SUPPORTING INFORMATION AVAILABLE

A description of the methods used to obtain the amide ^{15}N and ^{1}H assignments for CR1∼15,16^{-}, together with a table of chemical shift assignments. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Doolittle, R. F. (1985) *Trends Biochem. Sci. 114*, 233−237.
2. Bork, P., Shultz, J., and Ponting, C. P. (1997) *Trends Biochem Sci. 22*, 296−298.
3. Bork, P., Downing, A. K., Kieffer, B., and Campbell, I. D. (1995) *Q. Rev Biophys. 29*, 119−167.
4. Spitzfaden, C., Grant, R. P., Mardon, H., and Campbell, I. D. (1997) *J. Mol. Biol. 265*, 565−579.
5. Wiles, A. P., Shaw, G., Bright, J. Perczel, A. Campbell, I. D., and Barlow, P. N. (1997) *J. Mol. Biol. 272*, 253−265.
6. Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T., and Tack, B. F. (1986) *Immunol. Today 7*, 230−234.
7. Reid, K. B. M., and Day A. J. (1988) *Immunol. Today 10*, 177−180.
8. Norman, D. G., Barlow, P. N., Baron, M., Day, A. J., Sim, R. B., and Campbell, I. D. (1991) *J. Mol. Biol. 219*, 717−725.
9. Barlow, P. N., Norman, D. G., Steinkasserer, A., Horne, T. J., Pearce, J., Sim, R. B., and Campbell, I. D. (1992) *Biochemistry 31*, 3626−3630.
10. 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.
11. Krych, M., Hourcade, D., and Atkinson, J. P. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4353−4357.
12. Krych, M., Clemenza, L., Howdeshell, D., Hauhart, R., Horcade, D., and Atkinson, J. P. (1994) *J. Biol. Chem. 18*, 13273−13278.
13. Subramanian, V. B., Clemenza, L., Krych, M., and Atkinson J. P. (1996) *J. Immunol. 157*, 1242−1247.
14. Krych, M., Hauhart, R., and Atkinson, J. P. (1998) *J. Biol. Chem. 273*, 8623−8629.
15. Martin, D. R., Kalli, K. R., Yuryev, A., Fearon D., and Ahearn, J. M. (1991) *J. Exp. Med 174*, 1299−1311.
16. Weisman, H. F., Bartow, T., Leppo, M. K., Marsh, H. C., Carson, G. R., Concino, M. F., Boyle, M. P., Roux, K. H., Weisfeldt, M. L., and Fearon, D. T. (1990) *Science 249*, 146−151.
17. Hebell, T., Ahearn, J. M., and Fearon, D. T. (1991) *Science 254*, 102−105.
18. Kalli, K. R., Hsu, P., and Fearon, D. T. (1994) *Springer Semin. Immunopathol. 15*, 417−431.
19. Dorling, A., Riesbeck, K., Warrens, A., and Lechler, R. (1997) *Lancet 349*, 867−871.
20. Lowell, C. A., Klickstein, L. B., Carter, R. H., Mitchell, J. A., Fearon, D. T., and Ahearn, J. M. (1989) *J. Exp. Med. 170*, 1931−1943.
21. Ward, T, Pipkin, P. A., Clarkson, N. A., Stone, D. M., Minor, P. D., and Almond, J. W. (1994) *EMBO J. 13*, 5070−5074.
22. Dorig, R. E., Marcil, A., Chopra, A., and Richardson C. D. (1993) *Cell 75*, 295−305.
23. Klickstein, L. B., Wong, W. W., Smith, J. A., Weis, J. H., Wilson, J. G., and Fearon, D. T. (1987) *J. Exp. Med 165*, 1095−1112.
24. Krych, M., Hauhart, R., and Atkinson, J. P. (1996) *Mol. Immunol. 33*, 32.
25. Krych, M., Hauhart, R., Hourcade, D., Subramanian, B., and Atkinson, J. P. (1997) *J. Alergy Clin. Immunol. 99*, 1524.
26. Clark, N. S., Dodd, I, Mossakowska, D. E., Smith, R. A. G., and Gore, M. G. (1996) *Protein Eng. 9*, 877−884.
27. Scorer, C. A., Clare, J. J., McCombie, W. R. Romanos, M. A., and Sreekrishna, K. (1994) *Biotechnology 12*, 181−184.
28. O'Shea, J. J., Brown, E. J., Seligmann, B. E., Metcalf, J. A., Frank, M. M., and Gallin, J. I. (1985) *J. Immunol. 134*, 2580−2587.
29. Makrides, S. C., Scesney, S. M., Ford, P. J., Evans, K. S., Carson, G. R., and Marsh, H. C., Jr. (1992) *J. Biol. Chem. 267*, 24754−24761.
30. Privalov, P. L., and Potekhin S. A. (1986) *Methods Enzymol. 131*, 4 51.
31. Sober, H. A. (1970) *Handbook of Biochemistry*, 2nd ed., pp B75−76, CRC Press Inc. Boca Raton, FL.
32. Pace, C. N., and Schultz, J. M. (1990) *Protein Structure: a Practical Approach* (Creighton, T. E., Ed.) 2nd ed., pp 291−321, IRL Press.
33. Hamoguchi, K. (1992) *The Protein Molecule: Conformation Stability and Folding*, Springer-Verlag.
34. Colfen, H., and Harding, S. E. (1997) *Eur. Biophys J. Biophys Lett. 25*, 333−346.
35. Creeth J. M., and Harding, S. E. (1982) *J. Biochem. Biophys Methods 7*, 25−34.
36. Philo, S. J. (1997) *Biophys J. 72*, 435−444.
37. Harding, S. E. (1994) in *Methods in Molecular Biology 22, Optical spectroscopy and macroscopic techniques* (Jones, C., Mullow, A., and Thomas, A. H., Eds.) Humana Press, Towota, NJ.
38. Harding, S. E., and Colfen, H. (1995) *Anal. Biochem. 228*, 131−142.
39. Mori, S, Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. M. (1995) *J. Magn. Reson. B. 108*, 94−98.
40. van Geet, A. L. (1968) *Anal. Chem. 40*, 2227.
41. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
42. Kelly, S. M., and Price, N. C. (1997) *Biochim. Biophys. Acta 1338*, 161−185.
43. Hawrot, E., Xiao, Y., Shi, Q.-L, Norman, D., Kirkitadze, M., and Barlow, P. N. *FEBS Lett. 432*, 103−108.
44. Freskgård, P. O., Mårtensson, L. G., Johansson, P., Jonsson, B. H., and Carlsson U. (1994) *Biochemistry 33*, 14281−14288.
45. Di Scippio, R. G. (1992) *J. Immunol. 149*, 2592−2599.
46. Moore, M. D., Di Scippio, R. G., Cooper, N. R., and Nemerrow, G. R. (1989) *J. Biol. Chem. 264*, 20576−20582.
47. Aslam, M., Guthridge, J. M., Quigg, R. J., Holers, V. M., and Perkins, S. J. (1998) *Mol. Immunol. 35*, 389.
48. Medved, L. V., Busby, T. F., and Ingham, K. C. (1989) *Biochemistry 28*, 5408−5414.

BI982453A