Micelle Stability: κ-Casein Structure and Function

LAWRENCE K. CREAMER,*[,1] JEFFREY E. PLOWMAN,† MICHAEL J. LIDDELL,‡
MARK H. SMITH,* and JEREMY P. HILL*
*Food Science Section, New Zealand Dairy Research Institute,
Palmerston North, New Zealand
†Wool Research Organisation of New Zealand,
Christchurch, New Zealand
‡Chemistry Department, James Cook University,
Cairns, Queensland, Australia

ABSTRACT

The stability of the casein micelle is dependent on the presence of κ-casein (CN) on the surface of the micelle where it functions as an interface between the hydrophobic caseins of the micelle interior and the aqueous environment. κ-Casein is also involved in thiol-catalyzed disulfide interchange reactions with the whey proteins during heat treatments and, after rennet cleavage, in the facilitation of micelle coagulation. These functions of κ-CN are regulated by the three-dimensional structure of the protein on the micelle surface. The usual means of determining structure are not available for κ-CN because this protein is strongly self-associating and has never been crystallized. Instead, algorithms were used to predict selected secondary structures and circular dichroism spectroscopy on κ-CN and the macropeptide released by chymosin. Three peptides were synthesized to cover the chymosin-sensitive site (His^{98}-Lys^{111}), the region in the macropeptide that could be helical (Pro^{130}-Ile^{153}), and the region between. Nuclear magnetic resonance spectroscopy showed that the peptide His^{98}-Lys^{111} was probably a β-strand with tight turns at each end. This hypothesis was confirmed by a study of the molecular dynamics showing that the C variant of κ-CN interacted less strongly with chymosin; consequently, the slow renneting time of milk that contains this protein was explainable. Both circular dichroism and nuclear magnetic resonance indicated that the peptide Pro^{130}-Ile^{153} was probably helical under normal physiological conditions. A preliminary study using nuclear magnetic resonance showed that the intervening peptide had no discernible secondary structure. Consequently, most of the β-sheet structure of κ-CN is likely in the para-κ-CN region.

(**Key words**: casein micelle, κ-casein structure, glycomacropeptide structure)

Abbreviation key: **BPTI** = bovine pancreatic trypsin inhibitor, **CD** = circular dichroism, **FTIR** = Fourier transform infrared spectroscopy, **NMR** = nuclear magnetic resonance, **NOE** = nuclear Overhauser effect (or enhancement), **TFE** = 2,2,2-trifluoroethanol, **2D** = two-dimensional.

INTRODUCTION

It is generally accepted that κ-CN resides at the surface of the casein micelle as it exists naturally (3, 14, 36). Although the other caseins do not seem to have a role that requires well-defined structures, κ-CN may well be more structured to fulfill its function as the interface between the calcium-sensitive caseins and milk serum. Individual κ-CN molecules have been speculated to crosslink into disulfide-bonded polymers with a structure such that the hydrophilic tails project into the milk serum and the hydrophobic regions attach to the micelle core. In addition to the natural function of κ-CN of keeping the micelle in suspension, κ-CN is readily hydrolyzed once milk is in the calf stomach, allowing the formation of a coagulum that can be readily digested. Chymosin, an enzyme from the calf abomasum, cleaves κ-CN very specifically to produce insoluble para-κ-CN, which remains at the surface of the micelle and causes the micelles to aggregate and form a clot. The soluble macropeptide portion of κ-CN remains in the serum and presumably becomes a target for further enzyme attack.

To perform these functions, κ-CN needs to bind to the calcium-casein complexes of the micelle core and must be able to prevent premature micellar coagulation. The chymosin cleavage site must be accessible to the enzyme and must be cleavable at about pH 6.5, at which pH the enzyme is not normally active. Early studies recognized that chymosin cleaved κ-CN at a Phe-Met bond to give a large peptide (known as the macropeptide, caseinomacropeptide, or glycomacro-

Received June 23, 1997.
Accepted March 16, 1998.
[1]Corresponding author. New Zealand Dairy Research Institute, Private Bag 11029, Dairy Farm Road, Palmerston North, New Zealand.

peptide) that was soluble in trichloroacetic acid and an extremely insoluble para-κ-CN. An early study (12) using nuclear magnetic resonance (**NMR**) indicated that the structure of the macropeptide portion was similar before and after cleavage.

Early studies (24) in structure prediction showed that, once the primary structure of the protein had been determined, 22.5% helix structures and 31% sheet structures were indicated for the protein, although the spectral analysis using circular dichroism (**CD**) indicated 14% helix structures and 31% sheet structures, respectively (24). More recent predictions (23) suggest 10% helix and 30% sheet structures. Some CD work (11, 25) has indicated that the helical structures are more likely to be in the macropeptide region of the protein and that the sheet structures are more likely to be in the para-κ-CN region.

In 1989, when we first seriously contemplated determining the tertiary structure of this protein in solution or in its native state (attached to the micelle surface), we considered a number of possible approaches.

1. Structure prediction. Some preliminary work had been done (5) using this approach, but we did not have the tools for a more sophisticated approach.
2. Secondary structure assessment using CD, Fourier transform infrared spectroscopy (**FTIR**), or both. Although this assessment had the potential to give more information, it was not a particularly high resolution method.
3. X-Ray crystallography. Unfortunately, this option was not viable because neither κ-CN nor any of its fragments had been crystallized.
4. High resolution NMR. This technique appeared to have considerable potential and had the ability to show the position in space of every atom of the protein, provided that the appropriate spectra could be obtained and interpreted.
5. Deuterium exchange techniques were also considered to be useful for examining the effect of environmental changes on structure, but those techniques need to be used in conjunction with CD, FTIR, or ^{1}H (proton) NMR.

We chose to examine a series of relatively small peptides from κ-CN by high resolution proton NMR in conjunction with CD. This approach has since been applied to the N-terminal sequence of β-CN where nascent helical structures were observed (33, 35). In addition, restrained molecular dynamics techniques were applied to the peptide encompassing the bovine chymosin-cleavable bond (Phe^{105}-Met^{106}) as it is positioned in the active site cleft of chymosin.

PEPTIDE STRUCTURE AND THE ^{1}H NMR APPROACH

For peptides in a random conformation (i.e., when the whole peptide has high intramolecular mobility and the space available to every proton is constrained by the covalent bonds alone), no proton will be close to any other particular proton long enough to give a substantial ^{1}HMR nuclear Overhauser effect (**NOE**) (through-space) crosspeaks (other than αN i,i + 1). If the environment around this highly flexible peptide alters to change the energetics of peptides and solvents, then some regions in torsional space become less accessible, and certain peptide folds will appear to be relatively stable (i.e., the occupancy of certain conformational states, such as helical, sheet, or reverse-turn, increases). Under these circumstances, two-dimensional (**2D**) NMR NOE crosspeaks appear and have an intensity that is approximately proportional to $1/r^6$, where r is the internuclear distance and, to a lesser extent, the length of time that this particular conformational state is occupied. It is not unusual to detect several different conformations (e.g., α-helix and 3-10 helix) for the same residues by NMR. In theory, every conformation with an appropriate lifetime will be observed using the NMR technique, but, in practice, only those conformations with a moderate (more than about 20%) occupancy give peaks that are sufficiently above the noise level in the 2D spectra. (In contrast to the long time scale of the NMR experiment, the CD time scale is quite short, and more transient conformational states can be detected. Thus, the two techniques may give seemingly different structural information.)

Prior to the use of NOE spectra to provide structural information, the spectral peaks of the chemical shift have to be assigned uniquely and unambiguously to particular protons within the peptide, which is the single biggest task in the NMR analysis procedure.

Some Potential Difficulties with the Small Peptide Approach

Analysis of the structures of a series of small peptides to determine the structure of a protein has some inherent problems. In general, the conformation of a particular amino acid sequence within a protein in buffer solution is different from that of the isolated peptide sequence (or any of a series of similar se-

quences) in the same solution. This difference is less likely to occur in peptides that are essentially random coil in structure (in the protein) or that encompass a complete helical motif (2). Examples of β-sheet formation in isolated peptides are rare because the component strands in any β-sheet can come from different segments of the protein sequence, and the likelihood of two or more of these being present in any such peptide is low (6, 7). However, peptides that mimic natural β-sheet segments and that fold into β-sheet structures in aqueous solution have been designed (18). Peptides that are part of a β-sheet structure in the protein and have a high potential to form a β-sheet can exist as extended structure in solution.

A second problem that can arise is the low solubility of many peptides, which can be overcome by using mixed solvent systems. Therefore, the results may need interpretation for relevance to the structure of the protein under particular (e.g., physiological) conditions.

A third problem that can arise is that, if several amino acids in the sequence are of the same type (e.g., Glu) and in the same type of magnetic environment, then the proton signals may be indistinguishable and, thus, assignment becomes more difficult. This problem is an acute one for studies involving caseins, because there is a significant amount of random structure and a preponderance of certain amino acids, especially Thr in the κ-CN macropeptide and Pro in β-CN. Proline also interferes with the assignment process because there is no proton on the imino nitrogen atom, but all other amino acids have a single proton on the peptide bond nitrogen. Some of these latter problems can be overcome by appropriate choices of the peptide sequences being examined.

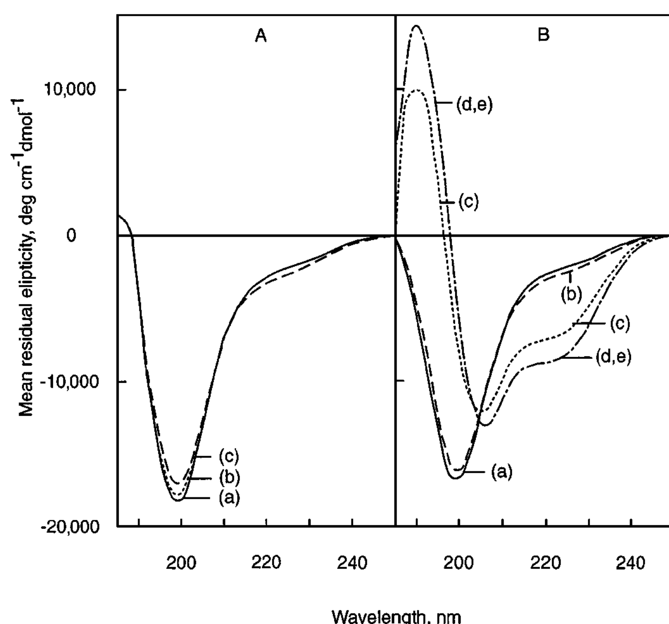

Figure 1. Circular dichroism spectrum of κ-CN macropeptide. A. Peptide dissolved in 20 m*M* NaCl solution: a, pH 7.2; b, pH 5.1; and c, pH 3.55. B. Peptide dissolved in 20 m*M* NaCl and 10 m*M* SDS solution: a, pH 7.05; b, pH 5.8; c, pH 5.0; d, pH 3.9; and e, pH 3.18. The peptide concentration was 0.12 mg/ml, and the spectra were obtained and analyzed as described previously (27).

Figure 2. Peptide sequences selected for examination by circular dichroism and proton nuclear magnetic resonance spectroscopy. The one-letter codes and the corresponding three-letter amino acid codes are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr; Z, Ser-PO_4 (not conventional).

Peptide Selection

Examination by CD (Figure 1) showed that the macropeptide contained about 25% α-helix, which corresponds to about 16 amino acids (or 9.5% of the 169 amino acids of κ-CN). Because the α-helix content of κ-CN is considered to be 10 to 15% (23), most of this helix must reside in the macropeptide. This result supports the earlier CD findings (11, 25). Most predictions put the macropeptide helix between residues 135 and 150 (22, 23, 27).

Another region of considerable interest is the cleavage site for chymosin. The sequences of the peptides that were chosen for our study are shown in Figure 2 and are labeled A to D. Peptide A (His^{98}-Lys^{111}) covers the chymosin-sensitive site, peptide B (Pro^{130}-Ile^{153}) encompasses the predicted helix, and peptides C (Met^{106}-Thr^{131}) and D (Glu^{151}-Val^{169}) complete the series for the whole of the C-terminal region of the protein.

STRUCTURE OF PEPTIDE B: THE PUTATIVE HELIX

Structure prediction studies of several κ-CN indicated that this region was likely to be helical in bovine (22, 23, 24) κ-CN but not porcine, ovine, or caprine κ-CN (27).

In a solution containing more than 25% (wt/wt) of the structure-enhancing solvent 2,2,2-trifluoroethanol (**TFE**), CD spectra showed that 80% of the peptide

was in the α-helical conformation, as was the case when the peptide was in 10 mM cetyltrimethyl ammonium chloride at pH 6. In contrast, the CD spectra of the peptide in buffer or in 10 mM SDS were consistent with a low helical content. Acidification of these solutions to pH 2.85 resulted in a slight increase in the helical content of the peptide in buffer and more markedly in 10 mM SDS solution. When the peptide was in 5 mM $CaCl_2$ solution at neutral pH, the CD spectrum indicated that some ordered structure was present. Those results (27) indicated that Glu^{137}, Glu^{140}, Glu^{147}, and Glu^{151} are important in determining the stability of the putative helix.

Further examination of this peptide by NMR was in 40% (wt/wt) TFE buffer. The results from a number of 2D experiments with slightly different solution conditions (to obtain the best resolution of the various proton peaks) were used to assign the resonances. Initially the resonances of specific spin systems were tentatively assigned to particular amino acid types. Next was the sequence-specific assignment for the protons on the polypeptide backbone using the NOE crosspeaks between the Cα protons of one residue and the amide N protons of the following residue (i.e., αN i,i + 1 using the nomenclature of Wüthrich (37)). Once the assignments were completed, the NOE crosspeaks for other interactions were determined, and these are shown in Figure 3.

Figure 3. Schematic diagram summarizing the relative intensities of the nuclear Overhauser effect connectivities among protons of the backbone NH (or proline δCH) and the αCH for the 130–153 peptide. The data were collected as described in Plowman et al. (27), and the standard nomenclature of Wüthrich (37) is used. This figure was adapted from results of Plowman et al. (27).

Interpretation of NOE

The NOE effect arises from a dipolar interaction between two protons that are in close proximity to each another. This effect shows up as crosspeaks with an intensity that is proportional to the closeness of the contact and the occupancy of that particular conformational state in the appropriate 2D spectrum (27). The results for peptide B are summarized as shown in Figure 3. The length of each horizontal bar indicates the distance between the two interacting protons in terms of the number of amino acid residues; for example, NN i,i + 1 indicates that the backbone amide proton of residue i is close to the amide proton of residue i + 1. The thickness of the line indicates the intensity of the crosspeak. Particular conformations give rise to different patterns of NOE because of the relationships of the backbone protons to one another. β-Strand, or extended, structures have intense αN i,i + 1 peaks and moderate NN i,i + 1 peaks, and random structure gives rise to moderate intensity αN i,i + 1 peaks and no NN i,i + 1 peaks. Helical structures give moderately intense αN i,i + 1 peaks, moderately intense NN i,i + 1 peaks, and various NN and αN peaks for i,i + x, where x = 2, 3, or 4, depending on the types of helical conformation that is involved (37).

The NOE results shown in Figure 3 together with chemical shift results were interpreted (27) to indicate that the structure of peptide B (residues 130–153 of bovine κ-CN B) in 40% TFE was strongly α-helical between residues 136 and 149 with some occupancy of the 3-10 conformation from residues 137–140 and from 145–148. Residues 151 to 153 were in the extended conformation. The Thr-Pro and Ser-Pro bonds were considered to be in the *trans* conformation. The NMR studies indicate that 14 residues are likely to be involved in the putative helix, which is similar to the number of residues expected in the caseinomacropeptide based on the CD studies [Figure 1; (25)].

The helix showed some amphipathic character, and a helical wheel plot (Figure 4) indicated that Glu^{137}, Glu^{140}, and Glu^{147} would all be on one face of the helix. In such positions, they would be well situated to form salt bridges with counter-ions such as calcium, hence, explaining why this cation shifted the structural state of the peptide to the more ordered form and possibly why cetyltrimethyl ammonium chloride was more effective than SDS in stabilizing the helical structure (27).

Not many NOE crosspeaks were observable in neutral buffer solution, which suggested that the structure of the peptide at neutral pH in water was essentially in a random coil conformation.

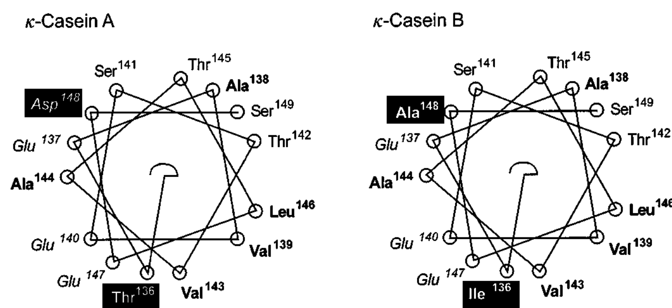

Figure 4. Helical wheel plots of the region from residues 136–149 in κ-CN A and B. Hydrophilic residues are shown as Roman type, hydrophobic residues as bold, and acidic residues as italic; the residues that differ between the A and B variants are highlighted. Potential sites of glycosylation are at Thr^{136} (A variant only), and Ser^{141} and the phosphorylation site is at Ser^{149}. This figure was adapted from results of Plowman et al. (27).

Effect of Environment on Peptide Structure

A number of recent studies have compared peptides under various conditions and their structures with the same amino acid segments within the protein environment. For example (13), three peptides that correspond to residues 11–28 (fragment 1), 61–77 (fragment 2), and 127–142 (fragment 3) of β-LG have been synthesized. In the protein, these fragments are mostly β-sheet, totally β-sheet, and totally α-helix, respectively (1), and have high, moderate, and high helical potential, respectively (4, 13, 17). Using CD analysis and the structure-inducing solvent TFE, fragment 3 was found to contain 10% helix when in water and over 50% helix when in 15% TFE, fragment 2 contained about 5% helix when in water and 15% helix when in 15% TFE, and fragment 1 contained 5% helix in water and about 30% helix when in 15% TFE. All of the fragments were 70 to 85% helix when in 50% TFE. A similar study by Kemmink and Creighton (21), using four peptides (1–15, 16–28, 29–44, and 45–58) that constitute the sequence of bovine pancreatic trypsin inhibitor (**BPTI**), found that these peptides achieved 13, 41, 23, and 43% helix by CD when dissolved in 50% TFE. However, NMR showed that no detectable effect of TFE on the structure of fragment 1 (f 1–15), although in BPTI this fragment contains several types of structure. Fragment 2 (f 16–28) was essentially helical by NMR and, in BPTI, consisted of β-sheet (f 18–24) followed by a reverse turn (f 25–28).

Clearly, an amino acid sequence that is helical within the context of the protein environment may not show a high helical occupancy in aqueous solution but could well do so in moderate concentrations of TFE. Conversely, an amino acid sequence that is essentially helical in 50% TFE need not be helical in the intact protein; the structure depends on the conformation of the rest of the protein (6, 7). The helicity determined by CD at 10 to 20% TFE can be indicative of the correct relationship between the structure in the peptide fragment and the parent protein, but NMR NOE crosspeaks at low levels of TFE are probably a better guide. In our case (27), 25% TFE was required to induce an 80% helical content at pH 3, which is a lower TFE concentration than expected for the induction of helicity in a peptide with a low to moderate natural propensity for the helical conformation.

The ability to attain a helical structure transiently could possibly be important for this region to act as the recipient of the posttranslational modifications of phosphorylation and glycosylation. The differences between the A and B variants (Figure 4) could well affect the structure of this helical region and influence glycosylation patterns.

PEPTIDE A: THE CHYMOSIN CLEAVAGE SITE

Because of the importance of the structure of the chymosin cleavage region, studies of NMR structure were initiated for appropriate peptides covering the cleavage regions of both the bovine and the human sequences (Figure 5). In addition, a study of restrained molecular modeling was initiated to compare the properties of chymosin and pepsin binding of the peptides covering this region of κ-CN.

NMR Structure of the Bovine Peptide

Predictive studies (5, 22, 23, 31) indicated that the structure of the peptide (His^{98}-Lys^{111}) from bovine κ-CN could be either β-sheet or α-helix, but was unlikely to be random structure.

The examination of the peptide A (Figure 2) encompassing the chymosin cleavage site (Phe^{105}-Met^{106}) by CD in TFE, in SDS, or in buffer alone at pH 6.7 indicated that sheet structure or helical secondary structure was negligible [i.e., no troughs in the 208 to 230-nm region of the far UV spectrum (J. E. Plowman, 1993, unpublished results)]. This result for the peptide in buffer alone did not preclude the presence of such structures in the intact protein, but the absence of such CD signals in TFE (for helix) and in SDS (for helix and sheet) suggested that such structures were unlikely.

```
VARRPRPHASF*IAIPPKKNQDKTAI Pig
 ARHPHPHLSF*MAIPPKKNQDKTE  Cow

 VRRPNLHPSF*IAIPPKKIQDKII  Human
 VPHPIPNPSF*LAIPTNEKHDNTA  Rat

  R P PH SF* AIPPKK QDKT   Consensus sequence
```

Figure 5. Comparative sequences of the chymosin sensitive region of cow, human, pig and rat κ-CN. The sequences are placed so that the cleaved bonds are aligned. Whenever the amino acid is identical for all four sequences the letter is shown in boldface in the consensus sequence. When three of the four amino acids are the same, the letter is shown in lightface. The one-letter codes and the corresponding three-letter amino acid codes are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr; Z, Ser-PO_4 (not conventional).

This peptide was only slightly soluble in aqueous solution and, consequently, was dissolved in hexadeuterated dimethylsulfoxide to achieve a suffiently high concentration of monomeric peptide (30). The NMR resonances for the peptide were assigned as just described. Preliminary experiments using NOE spectroscopy gave spectra without significant NOE crosspeaks. This result was considered to be related to the length and structure of the peptide; consequently, the peptide was then examined using rotating frame NOE spectroscopy (38). Although NOE crosspeaks were obtained, peaks were fewer and weaker than anticipated. The overall conclusion (30) was that the His-Pro and Ile-Pro bonds were in the *trans* conformation and that residues 102–108 were essentially in an extended conformation; Ile^{108} αC and N protons had unexpected chemical shifts, which was interpreted as an indication of slightly unusual structures near Ile^{108}.

Restrained Molecular Dynamics of Peptide A

Crystal structures were available for bovine chymosin, porcine pepsin, or rhizopuspepsin (with and without a bound inhibitor). The sequence and the three-dimensional structure of chymosin were compared with those of porcine pepsin and then with those of rhizopuspepsin to identify the structurally conserved regions. Then chymosin B and rhizopuspepsin plus inhibitor were spatially superimposed using computer graphics by aligning the residues close to the active site of each enzyme. The inhibitor was transferred from rhizopuspepsin to chymosin, and then the rhizopuspepsin structure was deleted, leaving the inhibitor in the chymosin cleft. The side chains along the inhibitor peptide were then sequentially changed into the appropriate casein side chains, and additional amino acid residues were added to each end of the peptide. On the basis of the NMR results (30), the added amino acids were initially put in the extended orientation. The energy of the structure was then minimized by allowing the peptide atoms to move within the fixed framework of the chymosin atoms. Then, the chymosin atoms within 5 Å of the peptide atoms were allowed to move, and the energy of the system was again minimized, which was followed by a series of steps in which the restraints on the system were lessened, the atoms were allowed to move, the restraints were reapplied, and the energy of the system was minimized (26).

The energy was lowest when the peptide was essentially in the extended conformation: Ile^{108} had a slightly unusual conformation, although the ϕ and ψ angles were within the normal range for β-strand structure (26). To allow His^{98} to be close to Asp^{247} of chymosin, the His^{98}-Pro^{99} bond was *cis* rather than *trans* [as indicated by the NMR study (30)].

It was recently noted that κ-CN C (19) and κ-CN G (8) are both cleaved by chymosin more slowly than are κ-CN A or B. The notable sequence difference is the substitution of Arg^{97} by His (19) and Cys (8), respectively. Possibly, Arg^{97} might have interacted with one or the other of the two Asp residues that were on the surface of the enzyme and have stabilized the initial peptide-enzyme complex. This scenario had not been considered previously because Arg^{97} was some distance from the cleavage site (Phe^{105}-Met^{106}).

Consequently, the procedure just described was applied to chymosin and to each of two elongated peptides: -Met-Ala-(His^{97} or Arg^{97})-His-Pro-His-Pro-His-Leu-Ser-Phe-Met-Ala-Ile-Pro-Pro-Lys, corresponding to κ-CN A or κ-CN B (Arg^{97}) and κ-CN C (His^{97}). If Arg^{97} was to interact with Asp^{249} of chymosin, then the His^{98}-Pro^{99} bond had to be *trans* (29, 32), as found by NMR (30), and not *cis*, as was suggested by the earlier study (26), so that His^{98} could interact with Asp^{247} of chymosin. The Arg^{97}-Asp^{249} of chymosin interaction was also sufficient to give greater stability to the complex between chymosin and peptide containing Arg than between chymosin and the peptide containing His.

NMR Structure of the Human Peptide

We have also examined the human equivalent to peptide A (Figure 2) encompassing the chymosin cleavage bond of human κ-CN by the NMR techniques developed for the 130–153 peptide (J. E. Plowman, L.

K. Creamer, M. J. Liddell, and J. J. Cross, 1998, unpublished). As with the bovine peptide, the NMR experiments were done in hexadeuterated dimethylsulfoxide for this study. The tetrapeptide sequence, Pro-Asn-Leu-His (Figure 5), was in a classical β-turn (J. E. Plowman, L. K. Creamer, M. J. Liddell, and J. J. Cross, 1998, unpublished) or was mostly in the extended structure, and the His-Pro bond was in the *trans* conformation.

Although a pseudo-gene for chymosin has been observed in humans, no functional neonatal protease that is similar to chymosin has been identified (9). Nevertheless, human κ-CN can be hydrolyzed by bovine chymosin, albeit at a reduced rate (10, 34) and, hence, provides another useful probe of the structural factors affecting chymosin activity. The reduced activity of chymosin for human κ-CN has been postulated to be due to the substitution of Met^{106} in the bovine protein for the bulkier side chain of the Ile residue in the human protein. However, the presence of a β-turn involving the residues Pro^{87}-His^{90} (Figure 5) would place Arg^{85} (Arg^{97} in the bovine protein) in a less favorable position for interaction with Asp^{249} in bovine chymosin. The structure and role of κ-CN in the human milk micelle are less clear, and the role of acid proteases in human milk coagulation and digestion needs to be clarified.

LIKELY OVERALL STRUCTURE OF κ-CN

At present our best estimate of κ-CN structure is summarized in Figure 6.

Macropeptide Region

The peptide Met^{106}-Thr^{131} (peptide C in Figure 2) was examined in aqueous solution buffered at both pH 3 and pH 5.2 on a 400-MHz spectrometer (J. E. Plowman and M. J. Liddell, 1994, unpublished results). Although good NOE spectra were obtained, only three NH i,i + 1 crosspeaks were observed, indicating that the peptide (and presumably this region in the intact protein) exists largely in the random coil conformation. This result was confirmed by David Craik (1994, personal communication), who examined our peptide using a 500-MHz spectrometer.

The 130–153 peptide (peptide B, Figure 2) was found to be helical in TFE concentrations above 25%, which means that this peptide is probably helical at times in the native protein. Unlike most proteins, κ-CN macropeptide residues are accessible to solvents (12), indicating that the macropeptide could be moving unhindered among various conformational states.

We were not able to synthesize peptide D (Figure 2), which was predicted to be 60% β-sheet and 40% random coil. Rather than pursue this matter, we

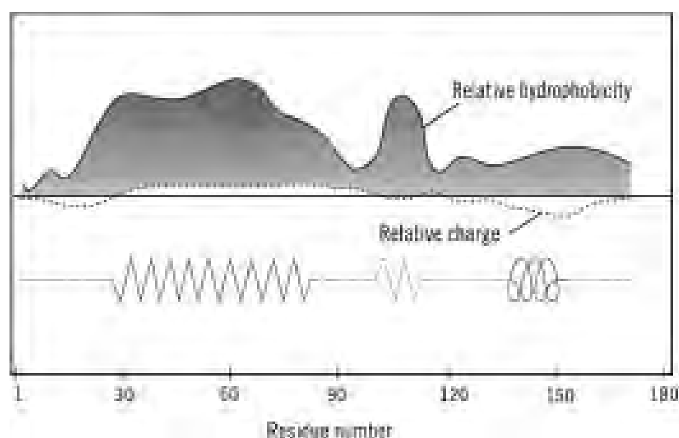

Figure 6. Likely secondary structures of various segments of bovine κ-CN. The upper curve and shaded area indicate the change in averaged hydrophobicity along the κ-CN sequence. The dashed line indicates the approximate net charge along the κ-CN sequence. The zigzag within the para-κ-CN region indicates the predicted β-sheet, the dashed zigzag indicates the β-sheet present when κ-CN is bound to chymosin, and the coil indicates the helix that is present under appropriate conditions.

decided that it would be more beneficial to examine the intact macropeptide, Met^{106}-Val^{169}, with a high field strength NMR spectrometer. Such a study is now underway (M. H. Smith and L. K. Creamer, 1998, unpublished).

Chymosin-Cleavable Region

The critical part of the segment from 95 to 111 (Figure 2) is a β-strand when in the active site cleft of chymosin and can adopt extended structures when not in this cleft. Possibly, Arg^{97} interacts with Asp^{249} of chymosin to help locate the sequence from 98 to 111 appropriately within the active site cleft prior to Phe^{105}-Met^{106} cleavage and release of the resultant fragments.

Para-κ-CN Region

From the N-terminus, the structure of the first segment of 21 amino acids is unclear. As has been pointed out previously (15, 16, 22, 23, 24, 27), this region is predicted to be helical despite a number of Pro residues, which makes extended stretches of helix unlikely. Because of the moderate number of ionizable residues, an NMR study of a synthetic peptide covering residues 1–21 of κ-CN would be feasible and could be helpful in clarifying the situation.

From residues 22 to 79, there is a high proportion of hydrophobic residues, no anionic residues, and few cationic residues. The region has been predicted to be β-sheet with intervening turns (22, 23, 24, 27), which

is probably the current best estimate of the likely secondary structure. As shown by CD, the overall secondary structure of κ-CN includes an amount of β-sheet (f 22–25), which is where most of this structure is likely to reside. It might be useful to examine this large peptide (f 22–79) by NMR in an environment that could mimic the micelle surface as it interacts with the κ-CN, for example, in SDS micelle solutions, which might mimic the hydrophobic environment of the casein micelle interior (20).

The sequence from residues 80–105 has not been unambiguously predicted to contain any particular structure. Because of the relatively low number and the even spread of hydrophobic residues and the high proportion of charged residues, a synthetic peptide covering this region of the protein could likely be examined successfully by NMR techniques.

CONCLUSIONS

Overall, the NMR technique has proved to be particularly useful for determining the most likely structure of κ-CN macropeptide within the natural environment (on the micelle surface), and the results are consistent with some earlier conclusions. Because of the nature of para-κ-CN (hydrophobic with few charged regions) and its close association with the other casein molecules in the micelle core, the likely structure of the whole native protein will be difficult to obtain. The sequences 1–21 and 80–105 (in the para-κ-CN region) could be examined by the techniques used so far, but, until more advanced NMR instrumentation and techniques can be brought to bear on the sequence 22–79, prediction methods and CD (or FTIR) remain the only available tools.

ACKNOWLEDGMENTS

We are grateful to Ted Baker at Massey University (Palmerston North, New Zealand) for assistance with the graphics equipment, to David Craik at the University of Queensland Cairns (Queensland, Australia) for assistance with the NMR studies, to Mike Boland for his support throughout this work, and to the Foundation for Research, Science, and Technology (Wellington, New Zealand) for Research Contracts DRI-402 and DRI-403.

REFERENCES

1 Brownlow, S., J. H. Morais Cabral, R. Cooper, D. R. Flower, S. J. Yewdall, I. Polikarpov, A.C.T. North, and L. Sawyer. 1997. Bovine β-lactoglobulin at 1.8Å resolution—still an enigmatic lipocalin. Structure 5:481–495.

2 Chakrabartty, A., and R. L. Baldwin. 1995. Stability of α-helices. Adv. Protein Chem. 46:141–176.

3 Creamer, L. K. 1991. Some aspects of casein micelle structure. Pages 148–163 *in* Interactions in Food Proteins. N. Parris and R. Barford, ed. ACS Symp. Ser. 454. Am. Chem. Soc., Washington, DC.

4 Creamer, L. K., D.A.D. Parry, and G. N. Malcolm. 1983. Secondary structure of bovine β-lactoglobulin B. Arch. Biochem. Biophys. 227:98–105.

5 Creamer, L. K., T. Richardson, and D.A.D. Parry. 1981. Secondary structure of bovine α_{S1}- and β-casein in solution. Arch. Biochem. Biophys. 211:689–696.

6 Dyson, H. J., and P. E. Wright. 1991. Defining solution conformations of small linear peptides. Ann. Rev. Biophys. Biophys. Chem. 20:519–538.

7 Dyson, H. J., and P. E. Wright. 1993. Peptide conformation and protein folding. Curr. Opin. Struct. Biol. 3:60–65.

8 Erhardt, G., E.-M. Prinzenberg, J. Buchberger, H. Krick-Saleck, I. Krause, and M. Müller. 1998 Bovine κ-casein G: detection, occurrence, molecular genetic characterisation, genotyping and coagulation properties. Pages 328–329 *in* Milk Protein Polymorphism. Int. Dairy Fed. Special Issue 9702. Int. Dairy Fed., Brussels, Belgium.

9 Foltmann, B. 1992. Chymosin: a short review on foetal and neonatal gastric proteases. Scand. J. Clin. Lab. Invest. 52: 65–79.

10 Green, M. R. 1986. Identification of human milk kappa-casein on polyacrylamide gels by differential staining with Ethyl-Stains-all and chymosin sensitivity. J. Histochem. Cytochem. 34:147–150.

11 Griffin, M.C.A., J. C. Price, and S. R. Martin. 1986. Effect of alcohols on the structure of caseins: circular dichroism studies of κ-casein A. Int. J. Biol. Macromol. 8:367–371.

12 Griffin, M.C.A., and G.C.K. Roberts. 1985. A ^{1}H-n.m.r. study of casein micelles. Biochem. J. 228:273–276.

13 Hamada, D., Y. Kuroda, T. Tanaka, and Y. Goto. 1995. High helical propensity of the peptide fragments derived from β-lactoglobulin, a predominantly β-sheet protein. J. Mol. Biol. 254:737–746.

14 Holt, C. 1992. Structure and stability of casein micelles. Adv. Prot. Chem. 43:63–151.

15 Holt, C., and L. Sawyer. 1988. Primary and predicted secondary structures of the caseins in relation to their biological functions. Prot. Eng. 2:251–259.

16 Holt, C., and L. Sawyer. 1993. Caseins as rheomorphic proteins: Interpretation of primary and secondary structures of the α_{s1}-, β- and κ-caseins. J. Chem. Soc. Faraday Trans. 89:2683–2692.

17 Ikeguchi, M., S. Kato, A. Shimizu, and S. Sugai. 1997. Molten globule state of equine β-lactoglobulin. Prot. Struct. Funct. Genet. 27:567–575.

18 Ilyina, E., V. Roongta, and K. H. Mayo. 1997. NMR structure of a *de novo* designed, peptide 33mer with two distinct, compact β-sheet folds. Biochemistry 36:5245–5250.

19 Jakob, E., C. Sievert, and Z. Puhan. 1994. Hydrolysis of κ-casein variants A, B and C by chymosin. Page 207 *in* Brief Commun. XXIV Int. Dairy Congr., Melbourne, 18–22 September 1994. Aust. Natl. Comm. Int. Dairy Fed., Glen Iris, Victoria, Australia.

20 Jarvet, J., J. Zdunek, P. Damberg, and A. Gräslund. 1997. Three-dimensional structure and position of porcine motilin in sodium dodecyl sulfate micelles determined by ^{1}H NMR. Biochemistry 36:8153–8163.

21 Kemmink, J., and T. E. Creighton. 1995. Effects of trifluoroethanol on the conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor. Biochemistry 34:12630–12635.

22 Kumosinski, T. F., E. M. Brown, and H. M. Farrell, Jr. 1991. Three-dimensional molecular modeling of bovine caseins: κ-casein. J. Dairy Sci. 74:2879–2887.

23 Kumosinski, T. F., E. M. Brown, and H. M. Farrell, Jr. 1993. Three-dimensional molecular modeling of bovine caseins: a refined, energy-minimized κ-casein structure. J. Dairy Sci. 76: 2507–2520.

24 Loucheux-Lefebvre, M.-H., J.-P. Aubert, and P. Jollès. 1978. Prediction of the conformation of the cow and sheep κ-caseins. Biophys. J. 23:323–336.
25 Ono, T., R. Yada, K. Yutani, and S. Nakai. 1987. Comparison of conformations of κ-casein, para-κ-casein and glycomacropeptide. Biochim. Biophys. Acta 911:318–325.
26 Plowman, J. E., and L. K. Creamer. 1995. Restrained molecular dynamics study of the interaction between bovine κ-casein peptide 98-111 and bovine chymosin and porcine pepsin. J. Dairy Res.62:451–467.
27 Plowman, J. E., L. K. Creamer, M. J. Liddell, and J. J. Cross. 1997. Solution conformation of a peptide corresponding to bovine κ-casein B residues 130-153 by circular dichroism spectroscopy and ^{1}H-nuclear magnetic resonance spectroscopy. J. Dairy Res. 64:377–397.
28 Deleted in proof.
29 Plowman, J. E., L. K. Creamer, M. H. Smith, and J. P. Hill. 1997. Restrained molecular dynamics investigation of the differences in association of chymosin to κ-caseins A and C. J. Dairy Res. 64:299–304.
30 Plowman, J. E., M. H. Smith, L. K. Creamer, M. J. Liddell, J. M. Coddington, J. J. Gibson, and D. R. Engelbretsen. 1994. Proton assignment and structural features of a peptide from the chymosin-sensitive region of bovine κ-casein determined by 2D-NMR spectroscopy. Magn. Reson. Chem. 32:458–464.
31 Raap, J., K.E.T. Kerling, H. J. Vreeman, and S. Visser. 1983. Peptide substrates for chymosin (rennin): conformational studies of κ-casein and some κ-casein-related oligopeptides by circular dichroism and secondary structure prediction. Arch. Biochem. Biophys. 221:117–124.
32 Smith, M. H., J. P. Hill, L. K. Creamer, and J. E. Plowman, 1998. Towards understanding the variant effect on the rate of cleavage by chymosin on κ-caseins A and C. Pages 185–188 *in* Milk Protein Polymorphism. Int. Dairy Fed. Special Issue 9702. Int. Dairy Fed., Brussels, Belgium.
33 Tsuda, S., R. Niki, T. Kuwata, I. Tanaka, and K. Hikichi. 1991. ^{1}H NMR study of casein phosphopeptide (1-25): assignment and conformation. Magn. Reson. Chem. 28:1097–1102.
34 Visser, S., P. J. van Rooijen, C. Schattenkerk, and K.E.T. Kerling. 1977. Peptide substrates for chymosin (rennin). Kinetic studies with bovine κ-casein-(103-108)-hexapeptide analogues. Biochim. Biophys. Acta 481:171–176.
35 Wahlgren, N. M., P. Djemek, and T. Drakenberg 1994. Secondary structures in β-casein peptide 1-42: a two dimensional nuclear magnetic resonance study. J. Dairy Res. 61:495–506.
36 Walstra, P. 1990. On the stability of casein micelles. J. Dairy Sci. 73:1965–1979.
37 Wüthrich, K. 1986. Pages 117–119 *in* NMR of Proteins and Nucleic Acids. John Wiley & Sons, New York, NY.
38 Zagorski, M. G. 1990. How to obtain accurate CAMELSPIN or ROESY spectra with identical low-power pulses for both the preparation and the mixing intervals. J. Magn. Reson. 89: 608–614.