

THREE-DIMENSIONAL MOLECULAR MODELING OF BOVINE CASEINS

5887

Abstract

Three-dimensional (3-D) structures derived from X-ray crystallography are important in elucidating structure-function relationships for many proteins. However, not all food proteins can be crystallized. The caseins of bovine milk are one class of non-crystallizable proteins (α_{s1} -, κ -, and β -). The complete primary and partial secondary structures of these proteins are known, but homologous proteins of known crystallographic structure cannot be found. Therefore, sequence based predictions of secondary structure were made and adjusted to conform with data from Raman and Fourier-transformed infra-red spectroscopy. With this information, 3-D structures for these caseins were built using the Sybyl molecular modeling programs. The κ -casein structure contained two anti-parallel β -sheets which are predominately hydrophobic. The α_{s1} -casein structure also contained a hydrophobic domain composed of β -sheets as well as a hydrophilic domain; these two are connected by a segment of α -helix. Both the κ - and α_{s1} -caseins represent unrefined models in that they have been manipulated to remove unrealistic bonds but have not been energy-minimized. Nevertheless the models account for the tendency of these caseins to associate. The β -casein model appears to follow a divergent structural pattern. When subjected to energy minimization, it yielded a loosely packed structure with an axial ratio of 2 to 1, a hydrophobic C-terminal domain, and a hydrophilic N-terminal end. All three casein structures showed good agreement with literature concerning their global biochemical and physico-chemical properties.

Key Words: Casein structure, protein functionality, milk proteins, three-dimensional molecular structure, molecular computer modeling, phosphorylation sites.

Introduction

The skim milk system contains a unique biocolloid, the casein micelle. This colloidal complex is in dynamic equilibrium with its environment. Changes in the state of the casein micelle system occur during milk secretion and processing (Farrell, 1988; Schmidt, 1982). Because of its innate importance to the milk system, the casein micelle has been studied extensively. From this large basis set of physical data, various investigators (for reviews see: Farrell, 1988; Schmidt, 1982) have attempted with varying degrees of success to formulate models which will adequately account for the properties of the casein micelle and allow for predictions of its behavior in real systems (Farrell, 1988). These models often fall short because of a serious gap in the data basis set. The combination of X-ray crystallography and molecular biology has contributed greatly to our understanding of the mechanisms of action of globular proteins. While the techniques of molecular biology are now being applied to casein, crystallographic structures will probably not be realized. It is this gap in our knowledge, the reconciliation of the physical chemical data for the caseins with molecular structure, that we have attempted to bridge with the tools of three-dimensional molecular modeling. These models are not an end in themselves; it is hoped that they will represent a new starting point in the examination and exploration of casein structure and function.

Materials and Methods

Predictions of secondary structures

Selection of appropriate conformational states for the individual amino acid residues was accomplished by comparing the results of sequence-based predictive techniques, primarily those of Chou and Fasman (1978), Garnier *et al.* (1978), and Cohen *et al.* (1983, 1986). Assignments of secondary structure (α -helix, β -sheet or β -turn) for the amino acid sequences were made when either predicted by more than one method, or strongly predicted by one and not predicted against by the others. In addition, because of the large number of proline

residues in the caseins, proline-based turn predictions were made using the data of Benedetti *et al.* (1983) and Ananthanarayanan *et al.* (1984).

Molecular modeling

The three-dimensional structures of the caseins were approximated using molecular modeling methods contained within the program Sybyl (Tripos, St. Louis, MO)¹ and implemented with an Evans and Sutherland (St. Louis, MO) PS390¹ interactive computer graphics display interfaced with a Silicon Graphics (Mountain View, CA) W-4D35 processor. In building the macromolecule we used a library or a dictionary of geometric parameters, i.e., bond lengths between specified atoms, bond angles, and van der Waals radii which are compatible with those found by X-ray crystallographic analysis. The molecular modeling software package contains this library or dictionary with average geometric parameters for each amino acid. Initial construction of the models followed the procedures previously developed (Kumosinski *et al.*, 1991a, 1991b, 1993). The construction of a model was accomplished by drawing upon the predicted behavior of the polypeptide chain from its amino acid sequence (see **Secondary structures** section above) and reconciling these predictions with spectroscopic data. Next, by examining the "domain" patterns arising from secondary structural predictions, combined with hydropathy plots (Farrell, 1988) peptide segments of from 60 to 100 residues were selected. Each peptide was built amino acid by amino acid, and assigned ϕ and ψ angles characteristic of the respective predicted structures for each residue. All ω angles were assigned the conventional trans configuration. In addition, aperiodic structures were in the extended rather than totally random configuration. This is a very important point, especially for the caseins which by spectroscopic and physical measurements do appear to have non-classical structures and are not "random walk" peptides. The choice of extended structure attempts to address this issue. The Sybyl subroutine "SCAN" was used, on the side chains only, to adjust torsional angles and relieve inappropriate van der Waals contacts. The individual pieces were then joined together to produce the total polypeptide model and readjusted with the "SCAN" subroutine. The models for α_{s1} - and κ -caseins were generated in this fashion. For β -casein further refinement was necessary.

Molecular force field-energy minimization

Quantum mechanical calculations have greatly facilitated the development of structure-function relationships for small molecules. However, many problems of biological interest, such as protein conformation, still require elementary model empirical energy functions. Though the functions are crude, this approach has been

applied successfully to the study of hydrocarbons, oligonucleotides, peptides, and proteins.

Here for β -casein we have employed a potential energy model which is simply described as a collection of overlapping balls for the atoms with given van der Waals radii connected by springs, and which mimic the vibrational character of the bonds. In addition, the atoms are assigned van der Waals attractive and repulsive forces, as well as electrostatic forces which quantify bonded and non-bonded interactions. Molecular mechanics or force field methods employ a combination of potential energy functions to optimize a structure. Three requirements for these force field calculations are an equation to calculate energy as a function of molecular geometry, parameterization (a set of best values for experimentally obtainable molecular properties), and an algorithm to calculate new atomic coordinates.

Empirical energy approaches are based on the assumption that one can replace a Born-Oppenheimer energy surface for a molecule, or system of molecules, by an analytical function. The potential energy function chosen is generally given as a sum of bond energies and non-bonded interaction terms (Weiner *et al.*, 1984, 1986). Three terms are used to represent the difference in energy between a geometry in which the bond lengths, bond angles, and dihedral angles have ideal values, and the actual geometry. The remaining terms represent non-bonded van der Waals and electrostatic interactions. The current version of AMBER by Kollman and coworkers (Weiner, *et al.*, 1986) includes the option of calculating hydrogen bond energies. This option was chosen for the β -casein work. The parameters used with this force field should include atomic partial charges calculated by Kollman using a united atom approach when only essential hydrogens are used. The calculated energy was minimized using a conjugate gradient algorithm.

Results and Discussion

Rationale for generation of secondary structure models

Various methodologies for sequence based secondary structural predictions are currently available, and these methods have been applied to the caseins (Creamer *et al.*, 1981; Loucheux-Lefebvre *et al.*, 1978; Raap *et al.*, 1983). In our initial approaches a sequence-based prediction was generated from each of three basis sets for the major proteins of milk: α_{s1} -, β - and κ -caseins. The results were ambiguous; for some portions of the sequences, structures (α -helix, etc.) were consistently predicted, while for others they were not. The relatively high proline contents of the caseins pose a problem since this residue, while occasionally found in both β -sheet and α -helical structures, is generally not favorable to either. In attempting to generate a consensus sequence-based prediction, we first focused upon solving the proline problem in a way commensurate with known behavior of this residue in proteins and model peptides.

Casein Molecular Modeling

Table 1. Comparison of adjusted sequence based predictions with spectroscopic data.

Sample		% Helix	% β -Structure	% Turns	% unspecified
α_{s1} -casein ^a	Raman ^b	8 - 13	18 - 20	29 - 35	33 - 40
	Initial	15	22	45	18
κ -Casein ^c	FTIR ^d	10	43	15	32
	Initial	16	27	37	20
β -casein ^e	Raman ^b	6 - 14	20 - 23	31 - 36	33 - 36
	Initial	8	18	40	39
	Final	10	20	34	34

^a α_{s1} -Casein is a single polypeptide chain of 199 residues.

^c κ -Casein is a single polypeptide chain of 169 residues.

^dCurley, Kumosinski and Farrell, preliminary data with 10% error.

^e β -Casein is a single polypeptide chain of 209 residues. Only this structure has been subjected to energy minimization and so it has an "initial" structure similar to the two others and a "final" structure following minimization.

The proline residues of the caseins are somewhat evenly dispersed throughout the sequences probably ruling out Type I and Type II polyproline structures suggested by Garnier (1966). It has been documented, however, that in synthetic peptides of known crystal structure, proline frequently occupies the second position of either a four-residue β -turn or a three-residue γ -turn (Benedetti *et al.*, 1983; Ananthanarayanan *et al.*, 1984; Rose *et al.*, 1985). Moreover, recent Raman and Fourier-transformed infra-red (FTIR) spectroscopic analyses of the caseins have suggested that by reference to known X-ray data and force constant calculations, up to 35% of the residues in caseins appear to be in β - or γ -turns (Byler *et al.*, 1988). To assess the probability that proline residues in the caseins might be located in turns, we examined the tetrapeptides containing proline in the second position for similarity with synthetic peptides known to form β - and γ -turns. These comparisons led to the hypothesis that most of the proline residues were in turns of some type. When the models were built, normal β -turns containing proline residues generally resulted in unfavorable van der Waals contacts with surrounding residues, leading us to assign some proline residues initially to the γ -turn conformation. β -Turns, other than those based on proline, were predicted by the Chou and Fasman (1978) and Cohen *et al.* (1986) methods. Many of the proline residues of β -casein, unlike those of α_{s1} - and κ -, are in Pro-Pro or Pro-X-Pro sequences. These sequences occur primarily in the C-terminal half of the molecule, thus these two special cases are important in β -casein. Using the method developed for κ -casein, where model peptides His-Pro-Pro-His and His-Pro-His-Pro-His were built and energy-minimized to test the best ϕ and ψ angles to be inserted in these sequences (Kumosinski *et al.*, 1991a), new models for β -casein, where -X- is hydrophobic, were constructed. These peptide models result in structures which do not unduly constrain the polypeptide chains and which have minimized

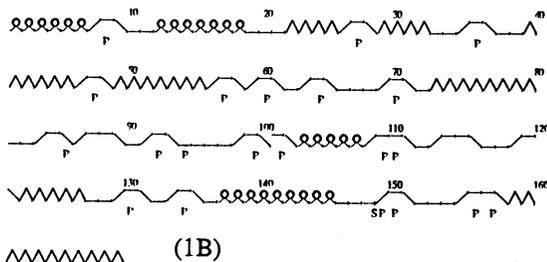
energies of approximately -8 kcal/residue, a value equivalent to that attained by energy minimization of X-ray crystallographic structures. For all of the above reasons, the total number of residues assigned to turn structures was only slightly in excess of the total predicted by spectroscopic methods (Table 1).

In a similar fashion, "consensus" scores for α -helix and β -sheet were arrived at by choosing from those regions having the highest predicted probability of a given structure to yield values in accord with the FTIR and Raman data (Byler *et al.*, 1988; and Curley, Kumosinski and Farrell, preliminary data with 10% error). In this case, all residues previously assigned to proline-based turns were eliminated first from consideration in α -helical or extended β -structures. A concern was that the assignment of all prolines to turns might decrease significantly the content of α -helix or β -sheet since proline can be the first residue in these structures. However, the net results of these calculations for β -, κ -, and α_{s1} -casein are compared with the spectroscopic data in Table 1 and the α -helical and β -structures are not overly reduced. Finally, all residues not included in these periodic structures were then considered to be in an extended aperiodic conformation. The initial conformational assignments for κ -casein are shown with its sequence in Figure 1, for α_{s1} -casein in Figure 2, and for β -casein in Figure 3. The κ -casein model is similar to that of Loucheux-Lefebvre *et al.* (1978) and with the exception of the β -caseins and γ -turns, the α_{s1} - and β -models are comparable to those of Creamer *et al.* (1981).

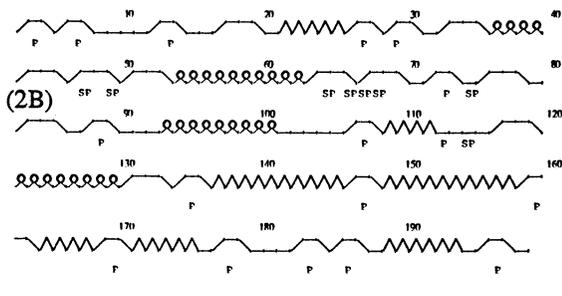
Rationale for generation of three-dimensional models

The secondary structural assignments which had been reconciled with FTIR and Raman spectroscopic data were used as a point of departure for the generation of three-dimensional models for α_{s1} -, β -, and κ -casein. Idealized ϕ and ψ angles assigned initially for each structural element are given in Table 2. The proteins were assembled as described in Materials and Methods.

OEONOEQPIRCEKDERFFSDKI AKYIPIQYVLSRYP SYGL
 NYTQOKPVVALINNOFLPYPYAKPAAVRS PAQILQWQVLS
 DTVPAKSCOAQPTTHARHFPHPHLSFHAIFPKKNQDKTEIF
 TINTIASGEP TSTPTIEAVESTVATLEASPEVIESPPEIN
 TVQVTSTAV (1A)



RPKHPKIKHQGLPOEVLNENLLRFFVAFPPQVFGKREKVNEL
 SKDIGSESTEGDAHEDIKQHEAESISSSEEIVPHNSVEQKH
 IQKEDVPSERYLGLYLEQLLRLLKRYKVPQLEIVPHNSAEERL
 HSHKEGIHAQOKEPHIGVNOELAYFYPELFRIFYQLDAYP
 SGANYVYVPLGTOYTDAPSFSDIPNFIGSENSEKTTHTPLW
 (2A)



RELEELNVPGEIVESLSSESSEESITRINKKIEKFOSEEOO
 TEDELQDKIHFFAQTOSLVYFFPGPIPNLSLQNIPLPTO
 PVVVPFFLOPEVHGVSKVKEAHAPKHKHEHFFPKYPVOFF
 ESQSLTLTDVENLHLLPFLLOSMMHQPFPPLPTVHFF
 SVLSLSQSKVLPVPEKAVYFPQRDHPJQAFLLYQORVLG
 VRCQFFPIIV (3A)

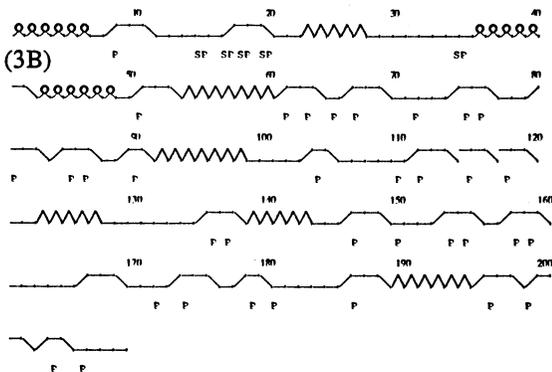


Table 2. Dihedral (ϕ, ψ) angles assigned to specific conformational states in the initial structures for the caseins.

Structures	Angle Assignments	
	ϕ	ψ
α -Helix ^a	-58	-47
β -Sheet ^a	-139	135
β -Turn ^a	N	-139
	N+1	-60
	N+2	-90
	N+3	-139
γ -Turn ^a	N	135
	N+1	-75
	N+2	81
X-Pro-Pro-X ^b	N	135
	N+1	-82
	N+2	-82
	N+3	81
Pro-X-Pro ^b	N	-82
	N+1	-72
	N+2	-82

^aDrawn from the literature, summarized in Rose *et al.* (1985).

^bKumosinski *et al.* (1993) average angles derived by molecular modeling for these residues in selected casein sequences.

Figure 1. (A) Sequence of κ -casein B; P below S denotes serine phosphate. (B) Summary of initial secondary structural assignments made for κ -casein; P denotes proline; SP denotes serine phosphate; α -helix denoted by scroll (e.g., residues 1-6); β -sheet by sawtooth (e.g., residues 40 to 46) and turns by turn-like structures (e.g., residues 35-37).

Figure 2. (A) Sequence of α_{s1} -casein B; P below S denotes serine phosphate. (B) Summary of initial secondary structural assignments made for α_{s1} -casein; P denotes proline; SP denotes serine phosphate; α -helix denoted by scroll (e.g., residues 37-40); β -sheet by sawtooth (e.g., residues 21 to 25) and turns by turn-like structures (e.g., residues 16-19).

Figure 3. (A) Sequence of β -casein A²; P below S denotes serine phosphate. (B) Summary of initial secondary structural assignments made for β -casein A²; P denotes proline; SP denotes serine phosphate; α -helix denoted by scroll (e.g., residues 1-6); β -sheet by sawtooth (e.g., residues 25 to 27) and turns by turn-like structures (e.g., residues 17-20).

Casein Molecular Modeling

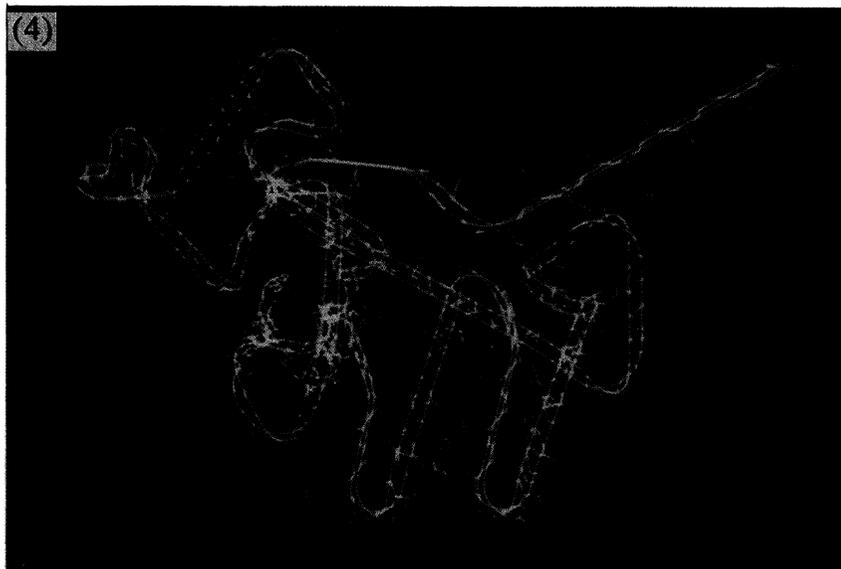


Figure 4. Three-dimensional molecular model of κ -casein. The peptide backbone is colored cyan, hydrophobic side chains green, acidic side chains red and basic side chains purple.



Figure 6. Three-dimensional molecular model of α_{s1} -casein. The peptide backbone is colored cyan, hydrophobic side chains green, acidic side chains red and basic side chains purple.



Figure 8. Three-dimensional model of β -casein-A². The hydrophobic side chains are colored green, acidic side chains are red and basic side chains purple. Side chain atoms of plasmin cleavage sites are colored red-orange, chymosin cleavage sites are orange and the C-terminal Val 209 is magenta. All other atoms are colored cyan and the backbone atoms are replaced by a double ribbon which traces the backbone and is colored yellow.

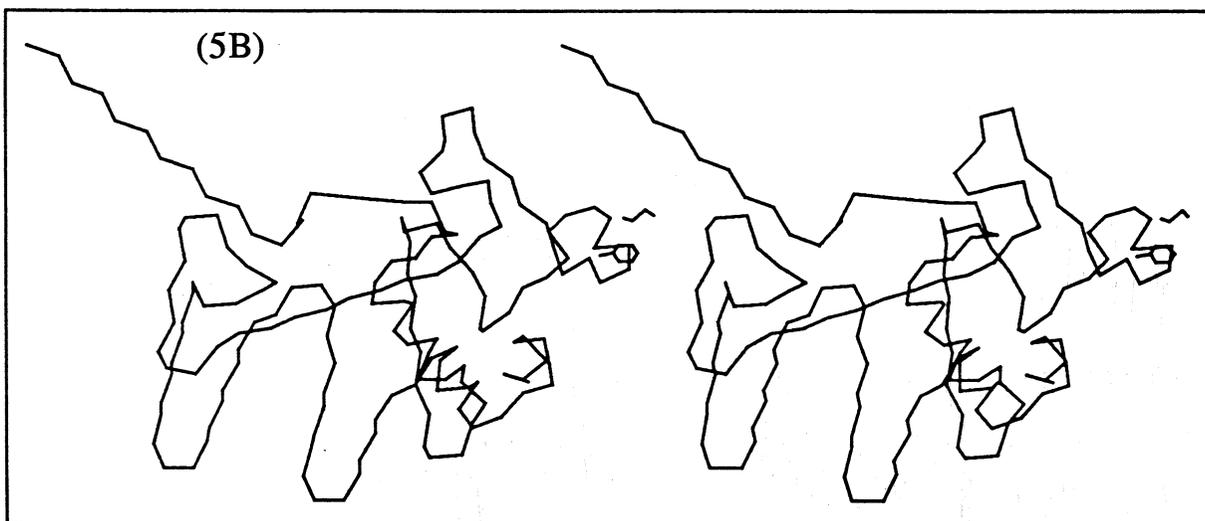
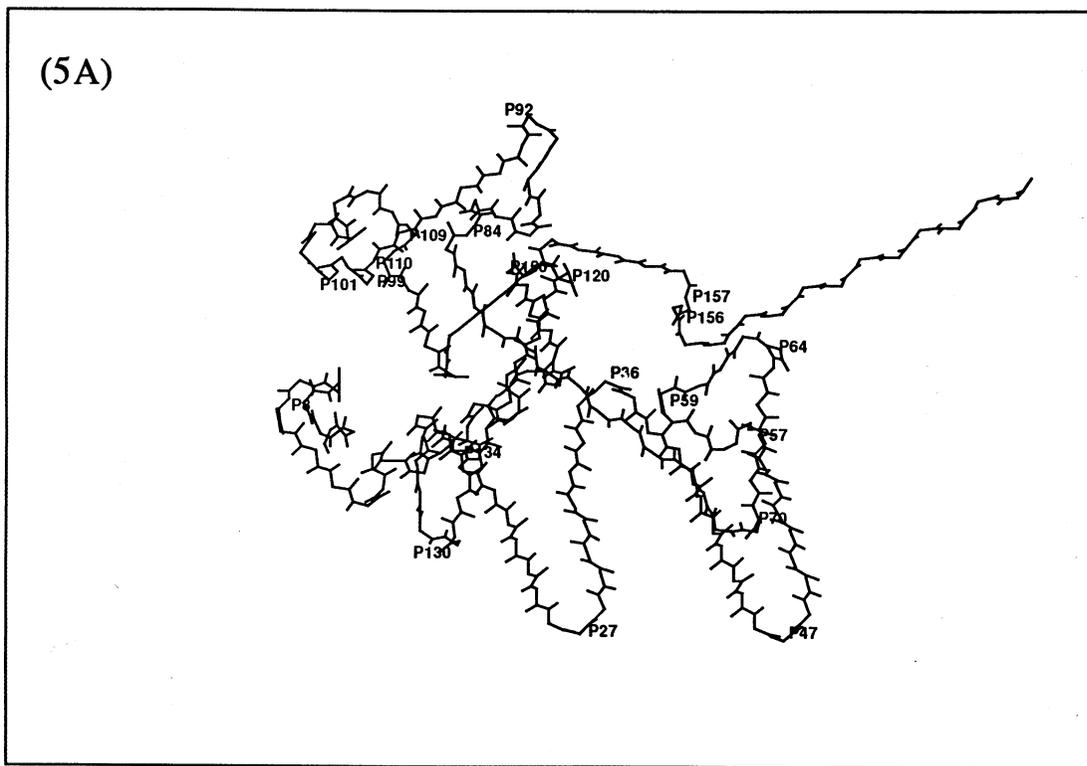


Figure 5. (A) Chain trace of κ -casein; prolines indicated. The model may be represented by a "horse and rider". The "horse" representing the legged structures and residues occupying the lower 2/3's of the figure, while the "rider" centers about Pro 92. (B) Stereo view of the three-dimensional molecular model of κ -casein, showing an α -carbon chain trace without side chains except phenylalanine-methionine 105-106 are labeled. Note: Fig. 6 is on color plate, p. 239.

Since the molecules were assembled from N \rightarrow C terminal, the procedure is analogous to the mechanism taking place *in vivo*, whereby the proteins are synthesized and presumably fold after insertion into the lumen of the endoplasmic reticulum (Farrell and Thompson, 1988).

Three-dimensional molecular model of κ -casein

The computer model generated as described above

for κ -casein is shown in Figure 4. Here, the hydrophobic side chains are colored green, serine phosphates, aspartic, and glutamic acid side chains are red, while the lysine, and arginine groups are colored purple. In descriptive terms, the protein can be thought of as being represented by a "horse and rider" as delineated in Figure 5A. Two distinct "legs" are seen in the "horse" portion of the model. These "legs" are generated by β -

Casein Molecular Modeling

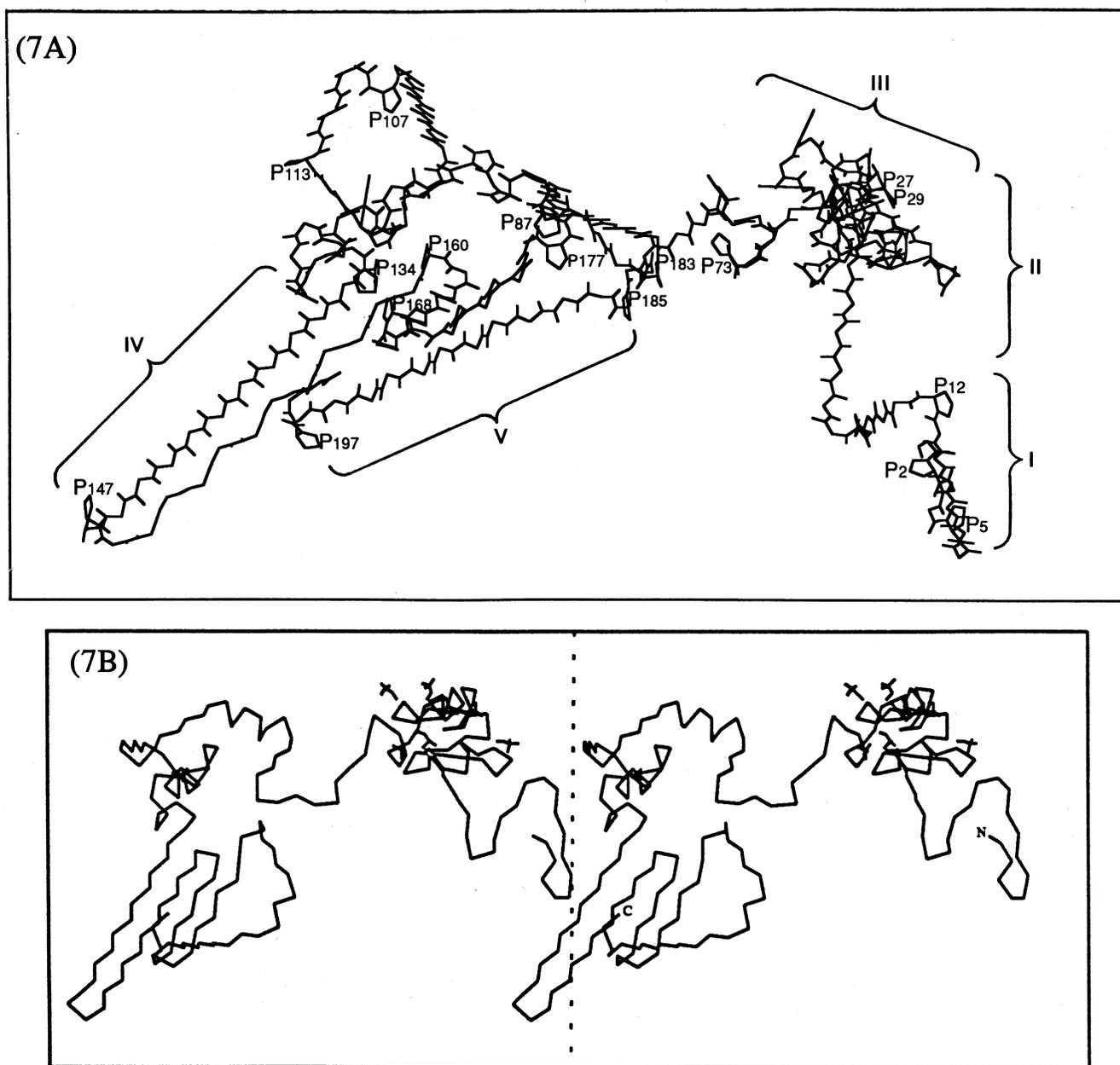


Figure 7. (A) Chain trace of α_{s1} -casein; prolines (P) indicated. Area I represents the hydrophilic N-terminal region; area II the α_{s1} -A deletion peptide; area III the major phosphopeptide region; area IV regions of hydrophobic β -sheet; and area V the C-terminal region. (B) Stereo view of the three-dimensional molecular model of α_{s1} -casein, showing an α -carbon chain trace without side chains except phosphoserines; the N- and C-terminal ends of the molecule are labelled. The ticked line represents the suggested stereo center, where the center of the stereo viewer should be placed.

sheet regions comprising residues 10 to 25; 29 to 34; 39 to 45 and 49 to 55, which are connected by γ - or β -turns. The overall dimensions of the κ -casein monomer predicted by this model are 8 x 6.2 x 5.8 nm. In the following section we will attempt to reconcile some known features of the chemistry of κ -casein with this molecular model. To aid in this discussion the back-

bone-only model is given in Figure 5A with prolines indicated, and in stereo in 5B.

Chemistry of κ -casein and the three-dimensional model

Chymosin: κ -Casein differs from α_s - and β -caseins in that it is soluble over a broad range of calcium ion

concentrations (Vaugh and Von Hippel, 1956). It was this calcium solubility which led these workers, upon discovering the κ -fraction, to assign to it the role of casein micelle stabilization. It is also the κ -casein fraction which is most readily cleaved by chymosin (rennin) (Jolles *et al.*, 1962, Kalan and Woychick, 1965); the resulting products are termed para- κ -casein and the macropeptide. It would appear that κ -casein is the key to micelle structure in that it stabilizes the calcium insoluble α_{s1} - and β -caseins.

The action of chymosin on the casein micelle is primarily the hydrolysis of the highly sensitive Phe-Met peptide bond (residues 105-106) of κ -casein. Sequence data (Mercier *et al.*, 1973) show this bond to be in a proline-rich region of the molecule between Pro-His-Pro (residues 99-101) and Pro-Pro (residues 109-110) which perhaps accounts for the high susceptibility of this specific bond to chymosin. From studies of model peptides, it has been suggested that the residues lying between Pro 101 and Pro 109 occur in a β -sheet structure (Raap *et al.*, 1983). Using other predictive methods, there is an equal chance that these residues could be in an α -helical conformation. In either case, the Pro-X-Pro and Pro-Pro residues cause the formation of a kink which neatly presents the otherwise hydrophobic Phe-Met on the surface of the molecule. In our model, we show the α -helix which represents the "horse's bit" in our descriptive "horse and rider" model. From model building considerations, this sequence represents the minimum number of residues for a stable α -helix. Cleavage of the Phe-Met (105-106) bond would render the helical conformation untenable, the helix would unwind and a considerable amount of configurational entropy would be added to the hydrolysis reaction. The change from β -sheet to extended conformation, though significant, would be lesser in its energy contribution. In either case, helix or sheet, the Pro-Pro and Pro-X-Pro turns present this segment on the surface and make it readily accessible to chymosin. All models for casein micelle structure must in some way account for this feature of κ -casein.

Hydrophobic interactions: Earlier speculation (Hill and Wake, 1969) that κ -casein might be a linear amphiphile seems to be only partially true. The amino-terminal fifth of the molecule has a relatively high charge frequency (0.34), however, the net charge is zero and this part of the protein is also relatively hydrophobic. Residues 35 to 68 represent an exceptionally hydrophobic area with almost no charge. It is precisely within this region that the majority of the residues found in the "legged" structures of the κ -casein molecule occur. These non-stranded, highly hydrophobic β -sheets make ideal sites for sheet-sheet interactions with other κ -casein molecules or with hydrophobic domains of α_s - and β -caseins. Indeed the concentration-dependent self-association profile of the reduced form of purified κ -casein can be fitted with a model for polymerization at a critical micelle concentration (Vreeman *et al.*, 1981). Several investigators (Rose, 1968; Downey and Murphy, 1970; Ali *et al.*, 1980; Davies and Law, 1983) have noted that

β - and κ -caseins diffuse out of the casein micelle at low temperatures. As one decreases the temperature, hydrophobic stabilization energy decreases, and κ -casein is able to dissociate from the micelle. Finally, the importance of the hydrophobic region in casein-casein interactions can be supported by the research of Woychik and Wondolowski (1969). Of nine tyrosine residues in κ -casein, seven are located between residues 35 and 68; nitration of 7 tyrosines in κ -casein severely inhibited its ability to stabilize α_{s1} -casein. As can be seen in Figure 4, the "legged" structures are constituted from this region and contain 7 tyrosine residues.

Sulfhydryl-disulfide interactions: κ -Casein contains two Cys residues: one resides at amino acid residue 11 and the other at amino acid residue 88. Whether these can form intra- or intermolecular disulfide bonds and the effects of such bonding on micelle stabilization have not been clearly established. The occurrence of free sulfhydryl groups in the milk-protein complex has been reported by Beeby (1964) but not by others (Swaigood *et al.*, 1964). The latter authors reported significant S-S cross-linking in purified κ -casein. However, Woychik *et al.* (1966) demonstrated that reduced and alkylated κ -casein stabilized α_{s1} -casein against calcium precipitation as well as native κ -casein. Pepper and Farrell (1982) found that in soluble whole casein, in the absence of Ca^{2+} , κ -casein occurred as a polydisperse high molecular weight complex. At low protein concentrations this complex could be separated from the other caseins by size exclusion gel chromatography in the absence of urea. The addition of reducing agents converted the κ -casein to the sulfhydryl-form which exhibited concentration-dependent associations, both with itself and with other caseins. Thus although κ -casein represents only 13% of the casein, many κ -casein molecules must be somewhat contiguous in the micelle in order to form these disulfide linked aggregates. This finding has been confirmed by Groves *et al.* (1992) who demonstrated that κ -casein occurs as a series of disulfide linked polymers ranging from monomer to octamer. It is interesting to note that Cys-11 is located between two segments predicted to be in α -helical conformations and is found on the "rider's" left, while Cys-88 is located in a predicted β -turn on the opposite side. In our model both of these residues are located near the surface of the molecule and are directed away from each other. This could account for the ability of the κ -casein molecule to form the inter-chain disulfide bonded polymers as discussed above, and quantitated by Rasmussen *et al.* (1992).

Sites for glycosylation and phosphorylation of κ -casein: Of the major components of the casein complex, only κ -casein can be glycosylated. Nearly all of the carbohydrate as well as the phosphate associated with κ -casein is bound to the macropeptide (Eigel *et al.*, 1984), which is the highly soluble portion released by chymosin hydrolysis. The major site for glycosylation, Thr-133, is found in our model on the back of the "horse" and is on a β -turn. The sites of phosphoryla-

tion, Ser-149 and Thr-145, are on the back portion and are found in β -turns as well.

Three-dimensional molecular model of α_{s1} -casein

The computer model, generated for α_{s1} -casein as described above is shown in Figure 6 (on color plate, p. 239) where it is displayed from carboxy- to amino-terminal (left to right). The molecule is colored as described for κ -casein. The best representation shows the molecule to be composed (right to left) of a hydrophilic amino-terminal portion, a segment of rather hydrophobic β -sheet, the phosphopeptide segment and then a very hydrophobic carboxy-terminal domain. For clarity the backbone without side chains is shown in Figure 7.

Chemistry of α_{s1} -casein and the three-dimensional model

Sites of phosphorylation in α_{s1} -casein: α_{s1} -Casein B is a single polypeptide chain with 199 amino acid residues and a molecular weight of 23,619 daltons (D) (Mercier *et al.*, 1971). The α_{s1} -B molecule contains eight phosphate residues, all in the form of serine monophosphates. Seven of these phosphoserine residues are clustered in an acidic portion of the molecule bounded by residues 43 and 80 (the second fifth of the molecule from the amino-terminal end). This highly acidic segment contains 12 carboxylic acid groups as well as the seven phosphoserines. The model shows the phosphoserine residues to be located on β -turns which is compatible with other known phosphorylated residues in proteins. This cluster forms a highly hydrophilic domain on the right shoulder of the molecule. This cluster is the major site for calcium binding, which in turn can be thermodynamically linked to the physico-chemical properties of α_{s1} -casein (Farrell *et al.*, 1988).

The α_{s1} -casein A deletion and chymosin cleavage sites: The rare α_{s1} -casein A genetic variant exhibits interactions which are highly temperature dependent. The A variant is the result of the sequential deletion of 13 amino acid residues bounded by residues 13 and 27; the majority of these deleted amino acids are apolar (Farrell *et al.*, 1988) but Glu 18 and 14 and Arg 22 are also deleted. The deleted segment encompasses a region predicted to be in a β -sheet. This sheet provides a spacer-arm between the hydrophilic amino-terminal region ("five o'clock" in Figure 6) and the phosphopeptide region. Deletion of this spacer-arm brings the hydrophilic section closer to the phosphate rich shoulder. In addition, the Phe-Phe bond (residues 23 and 24) is removed; this represents a major chymosin cleavage site (Mulvihill and Fox, 1979) and its absence may account for the poor quality products prepared from α_{s1} -casein A milks (Thompson *et al.*, 1969). Additional chymosin cleavage sites have been suggested (Mulvihill and Fox, 1979), but these do not appear to be relatively exposed for enzymatic attack.

Hydrophobic interactions: For α_{s1} -casein, the high degree of hydrophobicity exhibited by the segment containing residues 100 to 199 is probably responsible, in part, for the pronounced self-association of the

Table 3. Summary of the energy components (in kcal/mol) for energy-minimized β -casein A² model.

Bond Stretching Energy :	43.696
Angle Bending Energy :	627.030
Torsional Energy :	538.268
Out of Plane Bending Energy :	19.060
1-4 van der Waals Energy :	311.553
van der Waals Energy :	-1003.735
1-4 Electrostatic Energy :	2234.858
Electrostatic Energy :	-5454.868
H-Bond Energy :	-58.051
<hr/>	
Total Energy :	-2742.188
<hr/>	

α_{s1} -casein monomer in aqueous solution (Schmidt, 1982). This self-association approaches a limiting size under conditions of lowered ionic strength; the highly charged phosphopeptide region can readily account for this phenomenon through charge repulsions. It is noteworthy that at ionic strengths > 0.5 , α_{s1} -casein is salted out of solution at 37°C. This segment (residues 100 - 199) of α_{s1} -casein is highly enriched in hydrophobic residues and contains a segment of non-stranded β -sheet reminiscent of those found in κ -casein. This region is rich in tyrosine and as noted above for κ -casein, nitration of α_{s1} -casein also leads to decreased stability in reconstituted micellar structures (Woychik and Wondolowski, 1969).

β -Casein energy minimization of constructed model

The β -casein molecule is a single-chain polypeptide with five phosphoserine residues and a molecular weight of 23,980 D (Ribadeau-Dumas *et al.*, 1972). The high proline content of this molecule, relative to the other caseins, made construction of its model important for comparison with those developed above. The initial structure of β -casein was open and too diffuse, it was quite different from those of α_{s1} - and κ -casein (Kumosinski *et al.*, 1993). The model was, therefore, energy-minimized using the Kollman force field potential (Weiner *et al.*, 1986). Only essential hydrogen bonding protons were used in order to increase the speed of the calculation. Hence, van der Waals radii of carbon atoms were increased to account for the lack of hydrogens. This united atom approach is widely used when dealing with proteins in excess of thirty residues. Electrostatic interactions were added to the calculation by the use of united atom partial charges according to Weiner *et al.* (1986). Atoms at a distance of more than 0.8 nm (8 Å) were not considered as contributing to van der Waals and electrostatic interactions; the "non-bonded cut-off" used was thus 0.8 nm. The structure was minimized to a limit of ± 1 kcal and the results for each type of energy, as delineated in Equation 1, are presented in Table 3. Here, this total energy corresponds to a value of -13 kcal per residue which is consistent with values obtained

from energy minimization of structures derived from X-ray crystallography.

Energy-minimized molecular model of β -casein

The energy-minimized three-dimensional molecular model of β -casein A² is presented as a colored stick model in Figure 8 (on p. 239). The molecule is colored as described for κ -casein except for plasmin and chymosin cleavage sites (Fox, 1981; Visser, 1981); which are colored red-orange and orange, respectively; the backbone residues are traced by a double yellow ribbon. It is apparent from Figure 8 that although the structure of β -casein A² appears more compact than the previously derived structures for κ - and α_{s1} -caseins (Figures 4 and 6), it contains considerable asymmetric character. The backbone of the model demonstrates loops through which water can easily pass. The tertiary structure of β -casein is not that of a globular protein nor that of a solvent-denatured random coil. On inspection, the model can be thought to have a "crab-like" appearance with two large distorted hydrophilic arms. The lower one (at "four o'clock" position) results from multiple proline based turns (residues 85 to 119). The other large hydrophilic arm (residues 28 to 55) results from a combination of secondary structures which includes extended, sheet, turns and a segment of α -helix (residues 28 to 34); phosphoserine 35 occurs in this arm. Both arms contain relatively high charge frequency (Table 1) and the well characterized plasmin cleavage sites which occur at residues 28-29, 105-106 and 107-108.

Figure 8, shows that the overall shape of the molecule is asymmetric and can be approximated by a prolate ellipsoid with an axial ratio two to one. Furthermore, one end of the structure is predominately hydrophilic as noted above, and the rest consists mostly of a highly hydrophobic domain (left side of Figure 8). In β -casein, Pro-Pro and Pro-X-Pro sequences appear at intervals and result in highly convoluted hydrophobic segments or loops through which water may pass. The central X-residues of the Pro-X-Pro sequences are invariably non polar, as are 60% of residues flanking Pro-X-Pro and Pro-Pro units. These rigid hydrophobic segments may act as anchors to position portions of the sequence away from the surface (much as defined secondary structures can result in specific folding patterns for globular proteins) and thus to provide interaction sites for hydrophobically driven association reactions. In contrast, proline based γ -turns in the models of the α_{s1} - and κ -caseins result in hydrophobic, antiparallel β -sheets (Figures 4 and 6). This model for β -casein gives the impression of a large distorted surfactant molecule. This is most likely the result of the numerous turns introduced by the Pro-Pro and Pro-X-Pro residues. The kinds and types of proline based turns are given in Table 4. The surfactant nature of β -casein is underscored by a dipole moment of 1557 Debye and a net charge of -18.25 atomic charge units calculated from this structure using the Kollman (Weiner *et al.*, 1986) united atom partial charges; a value of only 815 Debye would be calculated from a spherically symmetric equivalent

Figure 9. (A) Chain trace of β -casein-A² with prolines (P) and serine phosphates (SP) indicated. (B) Stereo view, relaxed, of β -casein A² structure showing and α -carbon chain trace without side chains but N- and C-terminals are labeled. (C) Same orientation as in A; plasmin and chymosin cleavage sites are shown with residues labeled.

Table 4. Type of proline turn and sequence in β -casein.

β -turn ¹	γ -turn ²	Pro-Pro	Pro-X-Pro
9 - 10	70 - 72	75 - 76	61 - 67 ³
51 - 52	89 - 91	85 - 86	110 - 112
81 - 82	103 - 105	136 - 137	172 - 174
147 - 148	114 - 116	152 - 153	179 - 181
185 - 186	117 - 119	158 - 159	204 - 206
196 - 197			
200 - 201			

¹Middle amino acids of the four-residue β -turn.

²All three residues of the γ -turn.

³Residues in two successive Pro-X-Pro sequences resulting in a spring structure with the polypeptide chain proceeding in same direction.

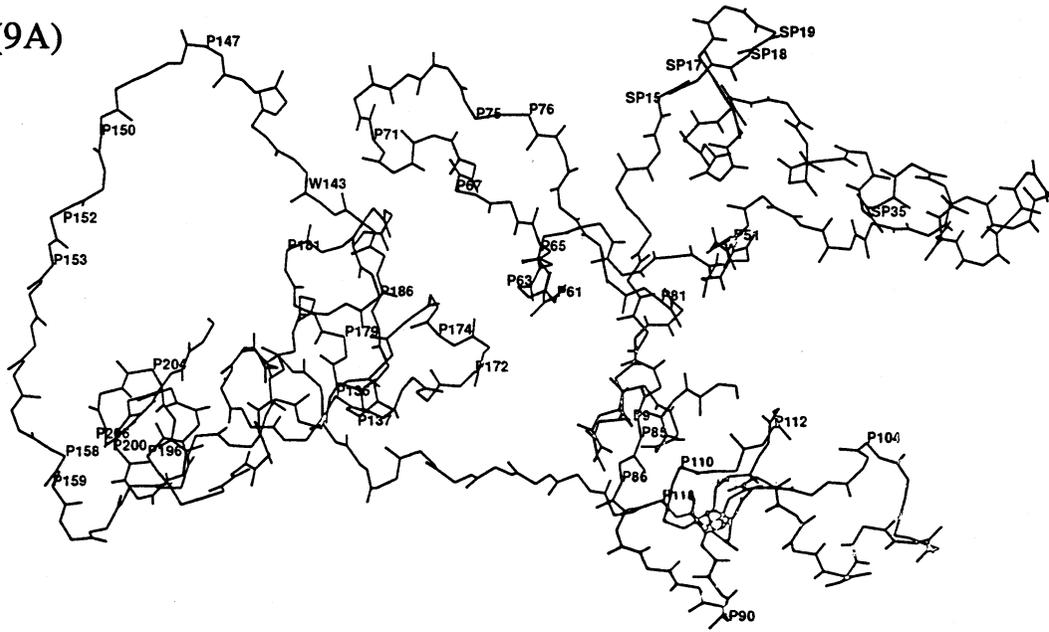
structure with the same net charge. A radius of gyration of 2.3 nm was calculated by assuming a solid prolate of revolution with axial ratio of 2 to 1. A backbone structure with labeled proline residues and a relaxed stereo view are also presented in Figures 9A and B. Labeled in Figure 9C are the plasmin and chymosin cleavage sites which all show accessibility for surface interactions with enzymes. Note that Trp-143 is relatively exposed in the monomer model. Pearce (1975) demonstrated that there is a blue shift of fluorescence when β -casein self-associates with increasing temperature. Thus Trp-143 would be exposed in the monomer, but buried in the polymer, accounting for the blue shift reported to occur on aggregation (Pearce, 1975). Also contained in the hydrophilic end of the structure are the phosphoserine residues which are located in a turn region as was the case for α_{s1} -casein (Figure 6). In fact, an extended turn region which resembles the crab's head (Figures 8, 9) ranges from residue 14 to 22 and contains 4 of the 5 phosphoserines in β -casein.

Secondary structure analysis

The energy-minimized three-dimensional structure of β -casein A² was compared with reported secondary structural analysis via Raman spectroscopy (Table 1). Figure 10 shows a Ramachandran plot of ϕ , ψ backbone dihedral angles (open circles) calculated from the refined β -casein A² structure using the Tripos' Sybyl molecular modeling software; areas of ϕ , ψ , angles which have been observed for known secondary structures (Rose *et al.*, 1985) are denoted. From this plot, the number of residues present, within the area assigned to a particular structure, can be estimated to obtain a rough value for

Casein Molecular Modeling

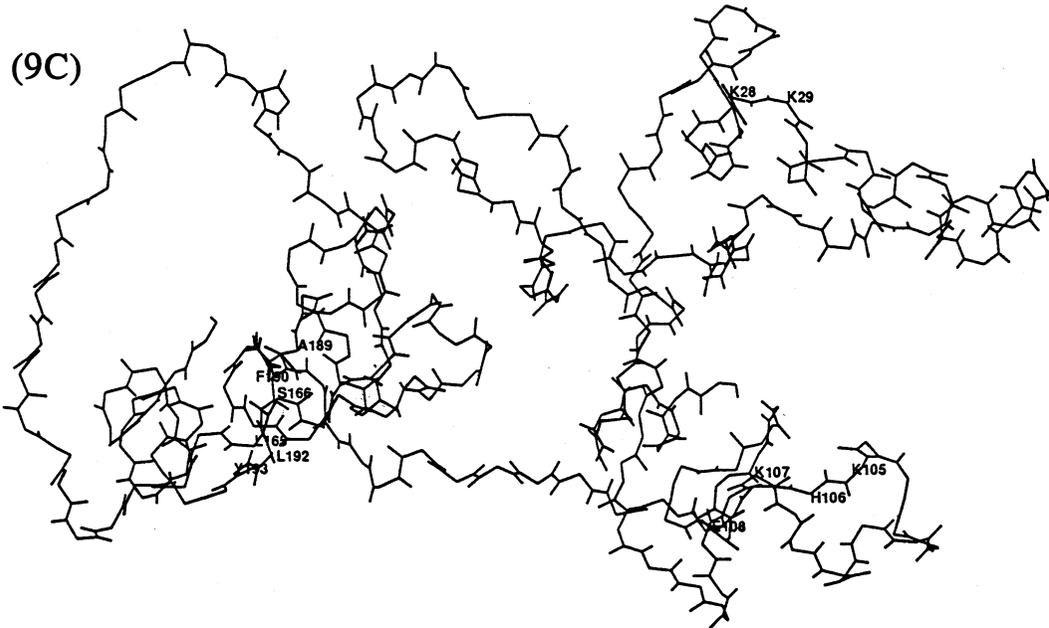
(9A)



(9B)



(9C)



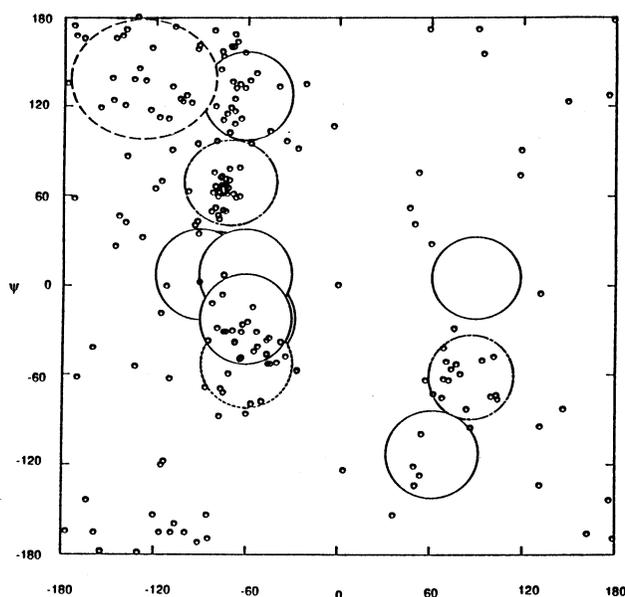


Figure 10. Ramachandran plot of ψ , ϕ angles calculated from energy-minimized structure of β -casein- A^2 . Dashed line (---) indicates area for β -sheet structure, dotted line (···) for α -helix, solid lines for β -turns, and dash-dot line (- · - · -) for γ -turns.

the global amounts of α -helix, β -sheet, or turns. However, this type of analysis alone can lead to overestimations since it does not take into account the required minimum residue length for a periodic structure (five to six residues). In addition, ϕ and ψ angles may not exactly represent the secondary structure due to changes in backbone bond lengths or bonding, so that visual inspection of the secondary structure by use of a chain trace or ribbon as in Figures 8 and 9 should also be employed in analyzing the secondary structure of a model.

The above procedure was used to analyze the β -casein A^2 structure; the global secondary structural results are in good agreement with those obtained in D_2O via Raman spectroscopy (Table 1). In addition to these classical structures, other types of non-classical structure are predicted. One type is the β -spiral arising from multiple proline turns as observed for residues 196-206 and 61-71 (Figures 9A and B). Another type of non-classical structure is the distorted hydrophilic arm discussed in the previous section, which also results from multiple proline turns contained within residues 85-90 and 100-120.

Plasmin and chymosin cleavage sites

Several investigators (for reviews see: Fox, 1981; Visser, 1981) have determined the sites of β -casein which are cleaved by the proteolytic enzymes plasmin and chymosin. It is of interest to further examine the β -casein A^2 three-dimensional model to ascertain whether these proteolytic sites are buried or exposed on its

surface. The principal plasmin cleavage sites, residues 28-29, 105-106 and 107-108, are colored red-orange on the model (Figure 8). All appear to be relatively exposed on the surface of the proposed structure. In fact, the cleavage sites are all located on the hydrophilic side of the molecule and particularly on the long distorted arms. Hence, all sites are not only solvent-accessible but also sufficiently exposed that in solution plasmin would readily hydrolyze these sites under the reported conditions of temperature and salt concentrations (Fox, 1981; Visser, 1981).

The chymosin cleavage sites are exposed on the monomer surface as well but are located on the hydrophobic (left) end of the proposed structure (see Figure 8). This model would predict that hydrophobic interactions, due to temperature-dependent self-association or interactions with other caseins, would inhibit chymosin action on β -casein. In fact, Visser (1981) and Creamer (1976) have shown that successful chymosin action on β -casein in solution occurs predominantly either at low temperatures where hydrophobically driven self-association is minimized or at lower pH. Berry and Creamer (1975) successfully limited hydrolysis to residues 189-190 on β -casein at 2°C; the resulting large polypeptide, i.e., residues 1-189, showed a marked decrease in self-association, relative to the parent β -casein by gel-permeation chromatography. In our model, residues 189-190 are located in a β -sheet structure (Figures 3 and 8), which is followed by a unique β -spiral (Figure 9A and B). This unique sheet-spiral structure is perpendicular to the hydrophobic surface of the model and is accessible to enzyme action in the monomer but not if hydrophobic self-associations occur.

Correlation of β -casein model with solution physico-chemical studies

The proposed three-dimensional structure for β -casein A^2 has a radius of gyration (R_g) of approximately 2.3 nm and dimensions of 2.1 by 4.2 nm, assuming a prolate ellipsoid of axial ratio two to one. Since the Sybyl molecular modeling software does not provide adequate surface (S) and volume (V) calculations, a V of 77.6 nm³ was estimated from the dimensions of the prolate ellipsoid model by assuming a S/V value of 2 nm⁻¹. This latter value was calculated by Kumosinski and Pessen (1982, 1985) for pepsin, which has a large surface area. Using this S/V ratio in conjunction with molecular weight and partial specific volume, a sedimentation constant of 1.7 S could be calculated for the β -casein. The actual value at 2°C is 1.5 S (Payens, 1979; Payens and van Markwijk, 1963). This calculated value of 1.7 S would be an upper limit and in reality a more precise surface area calculation might indeed yield a sedimentation coefficient much closer to the value of 1.5 S the experimentally determined by Payens and coworkers.

Andrews *et al.* (1979) studied the aggregation of β -casein by performing small-angle X-ray scattering experiments at 4°C and 25°C at protein concentrations ≥ 10 mg/ml. They calculated an R_g of 4.6 ± 0.2 nm for a presumed β -casein monomer at 4°C. However,

Casein Molecular Modeling

they did not estimate the value of the molecular weights for either the monomer or the observed polymer. Here, there is serious disagreement between their monomer Rg value of 4.6 nm and that determined from our structure of 2.3 nm. The value of 4.6 nm is extraordinarily high for protein of 24,000 D molecular weight and would be equivalent to a globular protein of molecular weight 312,000 D (Kumosinski and Pessen, 1982, 1985), or to completely denatured β -casein in a random coil with no secondary structure (Payens, 1979). To be consistent with an Rg of 4.6 nm and a sedimentation constant of 1.5 S, the monomer would have an axial ratio of 8 to 1 for the previously calculated volume of 77.6 nm³. The resulting ellipsoidal particle would have dimensions of 2.6 nm by 2.6 nm by 20.8 nm; even this hypothetical value represents a more compact structure than that of Payens and van Markwijk (1963) who calculated an axial ratio of 12 to 1 for β -casein at 4°C and Waugh *et al.* (1970) who calculated a 9 to 1 ratio. Noelken and Reibstein (1968) showed via sedimentation equilibrium studies that β -casein is monomeric at 2.5°C, but in the presence of ethylenediaminetetraacetic acid (EDTA) and at much lower concentration than the small-angle X-ray scattering experiments of Andrews *et al.* (1979). Waugh *et al.* (1970) suggested that reported differences in the physical chemical data of β -casein could be due to the presence or absence of Ca²⁺ resulting in some conformational changes. In the absence of consistent data, it would be prudent to assume that β -casein at 2°C is present, as a species intermediate between a totally random coil and a globular protein. Such a structure described in the above paragraph, a prolate ellipsoid of revolution with an axial ratio of two to one and a volume of 77.6 nm³, would indeed accommodate the intermediate type of structure suggested by Garnier (1966), Noelkin and Reibstein (1968), and Waugh *et al.* (1970). In fact, a globular protein of equal molecular weight (24,500 D) would have a volume of only 40 nm³ (Kumosinski and Pessen, 1982, 1985), which is significantly lower than 77.6 nm³ calculated for that of the refined β -casein A² model. This large volume is indicative of the relatively loose packing density which may characterize a structure intermediate between a random-walk polymer and a globular protein and which is also in accord with recent small-angle X-ray scattering results for whole casein (Pessen *et al.*, 1991).

Four explanations can be offered for the discrepancy between the experimentally observed Rg at 2°C and that of the model. The first, attributable initially to Garnier (1966), is that between 4° and 8°C a conformational change occurs which increases the degree of structure in β -casein. It is then this more structured molecule (perhaps similar to the model presented here) which participates in the aggregation reaction. The second, suggested by Pearce (1975), is that at 2°C β -casein is unstructured but as the self-association reaction occurs with increasing temperature, the monomer changes conformation in response to polymer formation (perhaps in a cooperative fashion) resulting in a more compact poly-

mer structure. A third concept drawn from Waugh *et al.* (1970) suggests that residual Ca²⁺ might cause conformational changes, leading to more compact structures. The fourth, offered here, is that the 4.6 nm Rg (Andrews *et al.*, 1979) represents, in fact, polymers of a more structured monomer, since the monomer molecular weight is only observed at very low concentrations with chelating agents in the cold. Evidence for a physical change of some sort in β -casein between 4 and 8°C has been presented by several investigators. Garnier (1966) observed changes in the ORD (optical rotatory dispersion) spectrum of bovine β -casein. Pearce (1975) observed marked native fluorescence changes between 2 and 10°C, and most recently Slattery and coworkers (Javor *et al.*, 1991) have observed changes in fluorescent dye binding and native fluorescence for human β -casein in this temperature region. The net result of all of these changes, regardless of the pathway, is that polymers of β -casein are more structured than β -casein monomers observed at 2°C and low concentrations in the presence of EDTA. The model presented here may thus represent either the kinetically active species in the polymerization reaction, or the final more compact β -casein monomer within its polymer.

Concluding Remarks

It must be stressed that the structures presented in this work represent preliminary models. They have been partially refined and poor van der Waals side chain contacts removed, but only the β -casein is an energy-minimized structure. For α_{s1} - and κ -casein electrostatic interactions, hydrogen bond formation, and backbone interactions were not taken into account. The refinement of these structures through the use of tools such as the Kollman Force Field (Weiner *et al.*, 1986) will be the thrust of future work. However, it is imperative to note that even these non-refined molecular modeling techniques (when combined with predicted secondary structures and spectroscopic results) can reveal important structure-function relationships.

For β -casein A², an energy-minimized three-dimensional model is constructed using a combination of secondary structure sequence-based prediction, global secondary structural results from Raman spectroscopy and molecular modeling techniques. This structure is in agreement with biochemical cleavage results for plasmin and chymosin action on β -casein. It is also in agreement with other experimentally derived results from sedimentation and small-angle X-ray scattering experiments and it provides a molecular basis for the temperature dependent self-association of β -casein.

The models of all three caseins are static and not dynamic and do not allow for a choice between the proposed mechanisms of conformational changes. However, these structures should be viewed as working models with the flexibility to be changed as more precise experiments are performed to ascertain the validity and predictability of this three-dimensional structure, which

of course is the nature of any scientific hypothesis. In future work, molecular dynamics calculations will be performed on these structures to test their stability when a kinetic energy equivalent to a bulk temperature is applied. However, the current models represent a starting point, and are assumed to represent an average dynamic structure. These structures will serve to allow comparisons with physico-chemical data and information from small-angle X-ray scattering. Finally all of this information can in turn be used in the genetic design of novel bovine proteins with altered functionality by the use of site-directed mutagenesis (Richardson *et al.*, 1992).

References

- Ali AE, Andrews AT, Cheeseman GC (1980). Influence of storage of milk on casein distribution between the micellar and soluble phases, and its relationship to cheese-making parameters. *J. Dairy Res.* **47**: 371-382.
- Ananthanarayanan VS, Brahmachari SK, Pattabiraman N (1984). Proline-containing β -turns in peptides and proteins: analysis of structural data on globular proteins. *Arch. Biochem. Biophys.* **232**: 482-495.
- Andrews AL, Atkinson D, Evans MTA, Finer EG, Green JP, Phillips MC, Robertson RN (1979). The conformation and aggregation of bovine β -casein A. I. Molecular aspects of thermal aggregation. *Biopolymers* **18**: 1105-1121.
- Beeby R (1964). The presence of sulfhydryl groups in κ -casein. *Biochim. Biophys. Acta.* **82**: 418-419.
- Benedetti E, Bavoso A, Blasio BD, Pavone V, Pedone C, Toniolo C, Bonora GM (1983). Solid-state geometry and conformation of linear, diastereoisomeric oligoprolines. *Biopolymers.* **22**: 305-317.
- Berry GP, Creamer LK (1975). The association of bovine β -casein. The importance of the C-terminal region. *Biochemistry* **14**: 3542-3545.
- Byler DM, Farrell HM Jr., Susi H (1988). Raman spectroscopic study of casein structure. *J. Dairy Sci.* **71**: 2622-2629.
- Chou PY, Fasman GD (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymology.* **47**: 45-148.
- Cohen FE, Abarbanel RM, Kuntz ID, Fletterick RJ (1983). Secondary structure assignment for α/β proteins by combinatorial approach. *Biochemistry.* **22**: 4894-4904.
- Cohen FE, Abarbanel RM, Kuntz ID, Fletterick RJ (1986). Turn prediction in proteins using a pattern-matching approach. *Biochemistry.* **25**: 266-275.
- Creamer LK (1976). A further study of the action of rennin on β -casein. *N.Z. J. Dairy Sci. and Technol.* **11**: 30-39.
- Creamer LK, Richardson T, Parry DAD (1981). Secondary structure of bovine α_{s1} - and β -casein in solution. *Arch. Biochem. Biophys.* **211**: 689-696.
- Davies DT, Law AJR (1983). Variation in the protein composition of bovine casein micelles and serum caseins in relation to micellar size and milk temperature. *J. Dairy Res.* **50**: 67-75.
- Downey WK, Murphy RF (1970). The temperature-dependent dissociation of β -casein from bovine casein micelles and complexes. *J. Dairy Res.* **37**: 361-372.
- Eigel WN, Butler JE, Ernstrom CA, Farrell HM Jr., Harwalkar VR, Jenness R, Whitney R McL (1984). Nomenclature of proteins of cows' milk: 5th revision. *J. Dairy Sci.* **67**: 1599-1631.
- Farrell HM Jr. (1988). Physical Equilibria: Proteins. In: *Fundamentals of Dairy Chemistry*, 3rd Edition. Wong N (ed.). Van Nostrand Reinhold, New York, pp. 461-510.
- Farrell HM Jr., Thompson MP (1988). The caseins of milk as calcium binding proteins. In: *Calcium binding proteins*. Thompson MP (ed.). CRC Press, Boca Raton, FL, pp. 118-137.
- Farrell HM Jr., Kumosinski TF, Pulaski P, Thompson MP (1988). Calcium induced associations of the caseins: A thermodynamic linkage approach to precipitation and resolubilization. *Arch. Biochem. Biophys.* **265**: 146-158.
- Fox PF (1981) Proteinases in dairy technology. *Neth. Milk Dairy J.* **35**: 233-253.
- Garnier J (1966) Conformation of β -casein in solution. Analysis of a thermal transition between 5° and 40°C. *J. Mol. Biol.* **19**: 586-590.
- Garnier J, Osguthorpe DJ, Robson B (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**: 97-120.
- Groves ML, Dower HJ, Farrell HM Jr. (1992). Reexamination of the polymeric distributions of κ -casein isolated from bovine milk. *J. Protein Chem.* **11**: 21-28.
- Hill RJ, Wake RG (1969). Amphiphile nature of κ -casein as the basis for its micelle stabilizing properties. *Nature.* **221**: 635-639.
- Javor GT, Sood SM, Chang P, Slattery CW (1991). Interactions of triply phosphorylated human β -casein: Fluorescence spectroscopy and light-scattering studies of conformation and self-association. *Arch. Biochem. Biophys.* **289**: 39-46.
- Jolles P, Alais C, Jolles J (1962). Amino acid composition of κ -casein and terminal amino acids of κ - and para κ -casein. *Arch. Biochem. Biophys.* **98**: 56-57.
- Kalan EB, Woychik JH (1965). Action of rennin on κ -casein, the amino acid compositions of the para κ -casein and glycomacropptide fractions. *J. Dairy Sci.* **48**: 1423-1428.
- Kumosinski TF, Pessen H (1982). Estimation of sedimentation coefficients of globular proteins: an application of small angle X-ray scattering. *Arch. Biochem. Biophys.* **219**: 89-100.
- Kumosinski TF, Pessen H (1985). Structural interpretation of hydrodynamic measurements of proteins in solution through correlations with X-ray data. *Methods Enzymol.* **117**: 154-182.

Casein Molecular Modeling

- Kumosinski TF, Brown EM, Farrell HM Jr. (1991a). Three dimensional molecular modeling of bovine caseins: κ -Casein. *J. Dairy Sci.* **74**: 2879-2887.
- Kumosinski TF, Brown EM, Farrell HM Jr. (1991b). Three dimensional molecular modeling of bovine caseins: α_{s1} -Casein. 1991. *J. Dairy Sci.* **74**: 2889-2895.
- Kumosinski TF, Brown EM, Farrell HM Jr. (1993). Three dimensional molecular modeling of bovine caseins: β -Casein. *J. Dairy Sci.* **76** (accepted for publication).
- Loucheux-Lefebvre M-H, Aubert J-P, Jolles P (1978). Prediction of the conformation of the cow and sheep κ -caseins. *Biophys. J.* **23**: 323-336.
- Mercier, JC, Grosclaude F, Ribadeau-Dumas B (1971). Primary structures of α_{s1} -casein bovine. Complete sequence. *Eur. J. Biochem.* **23**: 41-51.
- Mercier, JC, Brignon, G, Ribadeau-Dumas B (1973). Primary structure of κ -casein B bovine. Complete sequence. *Eur. J. Biochem.* **35**: 222-235.
- Mulvihill DM, Fox PF (1979). Proteolytic specificity of chymosin on bovine α_{s1} -casein. *J. Dairy Research.* **46**: 641-651.
- Noelken M, Reibstein M (1968). Conformation of β -casein B. *Arch. Biochem. Biophys.* **123**: 397-402.
- Payens TAJ (1979). Casein micelles: the colloidal-chemical approach. *J. Dairy Res.* **46**: 291-306.
- Payens TAJ, van Markwijk BW (1963). Some features of the association of β -casein. *Biochim. Biophys. Acta.* **71**: 517-530.
- Pearce KN (1975). A fluorescence study of the temperature dependent polymerization of bovine β -casein A¹. *Eur. J. Biochem.* **58**: 23-29.
- Pepper L, Farrell HM Jr. (1982). Interactions leading to formation of casein submicelles. *J. Dairy Sci.* **65**: 2259-2266.
- Pessen H, Kumosinski TF, Farrell HM Jr., Brumberger H (1991) Tertiary and quaternary structural differences between two genetic variants of bovine casein by small-angle X-ray scattering. *Arch. Biochem. Biophys.* **284**: 133-142.
- Raap J, Kerling KET, Vreeman HJ, Visser S (1983). Peptide substrates for chymosin (rennin): Conformational studies of κ -casein and some related κ -casein oligo-peptides by circular dichroism and secondary structural predictions. *Arch. Biochem. Biophys.* **221**: 117-124.
- Rasmussen, LK, Højrup P, Petersen TE (1992). The multimeric structure and disulfide-bonding pattern of bovine κ -casein. *Eur. J. Biochem.* **207**: 215-222.
- Ribadeau-Dumas B, Brignon G, Grosclaude F, Mercier JC (1972). Primary structure of β -casein bovine. Complete sequence. *Eur. J. Biochem.* **25**: 505-514.
- Richardson T, Oh S, Jiminez-Flores R, Kumosinski TF, Brown EM, Farrell HM Jr. (1992) Molecular modeling and genetic engineering of milk proteins. In: *Advanced Dairy Chemistry, 1: Proteins.* Fox PF (ed.). Elsevier, Essex, England, pp. 545-578.
- Rose D (1968). Relation between micellar and serum casein in bovine milk. *J. Dairy Sci.* **51**: 1897-1902.
- Rose GD, Gierasch LM, Smith JA (1985). Turns in peptides and proteins. *Adv. Prot. Chem.* **37**: 1-109.
- Schmidt DG (1982). Associations of caseins and casein micelle structure. In: *Developments in Dairy Chemistry -1.* Fox PF (ed.). Applied Science Pub. Ltd., Essex, England, pp. 61-82.
- Swaisgood HE, Brunner JR, Lillevik HA (1964). Physical parameters of κ -casein from cow's milk. *Biochemistry.* **3**: 1616-1623.
- Thompson MP, Gordon WG, Boswell RT, Farrell HM Jr. (1969). Solubility, solvation, and stabilization of α_{s1} - and β -caseins. *J. Dairy Sci.* **52**: 1166-1173.
- Visser S (1981). Proteolytic enzymes and their action on milk proteins. A review. *Neth. Milk. Dairy J.* **35**: 65-88.
- Vreeman HJ, Brinkhuis JA, Van Der Spek CA (1981). Some association properties of bovine SH- κ -casein. *Biophys. Chem.* **14**: 185-193.
- Waugh DF, Von Hippel PH (1956). κ -Casein and the stabilization of casein micelles. *J. Am. Chem. Soc.* **78**: 4576-4582.
- Waugh DF, Creamer LK, Slattery CW, Dresdner GW (1970). Core polymers of casein micelles. *Biochemistry* **9**: 786-795.
- Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner P (1984). A new force field for molecular simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**: 765-784.
- Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986). Force field calculations in computational chemistry. *J. Comput. Chem.* **7**: 230-252.
- Woychik JH, Wondolowski MV (1969). Nitration of tyrosyl residues in κ - and α_{s1} -caseins. *J. Dairy Sci.* **52**: 1669-1672.
- Woychik JH, Kalan EB, Noelken ME (1966). Chromatographic isolation and partial characterization of reduced κ -casein components. *Biochemistry.* **5**: 2276-2288.

Discussion with Reviewers

H.E. Swaisgood: What is the expected effect of omitting solvent water from these calculations?

Authors: As you are well aware, the caseins in general have a pronounced tendency to self-associate primarily, though not exclusively, through hydrophobic interactions. In the *in vacuo* models one of the "signatures" for α_{s1} - and κ -caseins is the proline-based hydrophobic turn. With neither water, nor a hydrophobic calculation, these hydrophobic areas appear on the monomer surface. These hydrophobic areas may thus serve as protein-protein interaction sites and in fact become buried in the polymer, which is in line with many experimental observations. In point of fact, the absence of these terms may be a positive element in these studies. For β -casein, the absence of water in the calculations may allow the molecule to contract somewhat more than it would as a

monomer, but most physical evidence suggests this occurs on association so that the net result is that the β -monomer could represent the monomer-within-the-polymer of associated β -caseins. In both cases, the monomers may, thus, ultimately serve as potential building-blocks for understanding casein micelle structure.

H.E. Swaisgood: How was the possibility of multiple energy minima handled?

Authors: In *de novo* calculations, the potential for false energy minima is great. However, in these studies we superimposed a number of constraints including secondary structural predictions, thus not every bond begins from a random start, and the most favored sections persist. The conjugate gradient algorithm used here circumvents part but not all of the problem. Molecular dynamics calculations are needed to further test the validity of the structures presented. These are currently underway and show good promise but the "final" answer is really not at hand. We feel that we are following a logical pattern to arrive at the best **working** model possible. Note the word "**working**" is emphasized.

M.N. Liebman: With respect to the conformation of β -casein in solution, could not the Ca^{2+} effects be monitored with FTIR to address its potential contribution or lack thereof?

Authors: That is a very good suggestion, and we have some preliminary data to suggest that for whole caseins there are subtle conformation changes with Ca^{2+} . Following the conformation of β -casein in a temperature-controlled FTIR cell is an excellent suggestion for further work.

M.N. Liebman: I would encourage the adoption of conservative analysis of the projected models and encourage their comparison with other proteins to try to identify structural motifs, components, characteristics which might be supportive of their segmental similarity of compliance with other aspects of protein structure.

Authors: These are **working** models and must to be altered and improved as more data are acquired. In essence, these models should represent the true scientific dialectic at work.

L.K. Creamer: With respect to the nitration of tyrosine and its effects on κ -casein, nitration as used by Woychik and Wondolowski was only employed for a short time until it was found that it commonly caused cross-links between tyrosine residues. Thus, these data need to be interpreted with caution. (In our hands, at about the same time, we obtained heavily cross-linked proteins).

Authors: This is an interesting observation because our thesis is that the hydrophobic areas containing tyrosines serve as sites for protein-protein interactions. If the cross-links are between κ -casein molecules, it serves to illustrate the point. If they are within a molecule, it would tell us something regarding the through-space distance of non-adjacent tyrosine residues in the molecule,

and these distances could be calculated from the software available and could help refine the model.

L.K. Creamer: One of the problems in all data-based, or statistical, prediction systems is how to deal with uncommon amino acid residues. It is not stated how the authors dealt with phosphoserine but from the result that phosphoserine was commonly found at β -bends, it seems likely that phosphoserine was treated as though it were serine. This is fine for predicting the structure of the unphosphorylated protein but would have drawbacks for the final protein with its high concentration of negative charges in very small volumes such as occur in α_{s1} - and β -casein in the absence of calcium.

Authors: The Sybyl programs allow construction of phosphoserine and this was done. In all the past and future work we use the phosphorylated form. In some recent trials we have also compared "dephosphorylated" forms and there are differences. The energy-minimized β -casein began and ended with phosphorylated serine residues found by sequence analysis.

L.K. Creamer: One of the major driving forces for proteins to adopt particular structures is the so-called hydrophobic effect (see for example in Dill, *Biochemistry* 29: 7133-7155). In the model-building described, the interactions of water with the protein, or the proline peptides was apparently neglected in the energy-minimization calculations. Will any of the future work include water interactions as a part of the calculations?

Authors: Addition of solvent water is a good idea for extension of this work as our ability to do protracted calculations will increase. A future need in molecular modeling for proteins and membranes would be the formulation of a reliable molecular modeling calculation for the hydrophobic effect. Such calculations are in their infancy at this time. As noted above in answer to Dr. Swaisgood's question, the absence of these calculations may not be a serious drawback, since our exposed hydrophobic surfaces on our monomers may ultimately be buried on polymerization, and the dominant physical feature of all caseins is their ability as purified proteins to self-associate or to enter into mixed associations with other caseins in micelle formation.