

Three-Dimensional Molecular Modeling of Bovine Caseins: An Energy-Minimized β -Casein Structure¹

T. F. KUMOSINSKI, E. M. BROWN, and H. M. FARRELL, JR.
Eastern Regional Research Center
Agricultural Research Service, USDA
600 E. Mermaid Lane
Philadelphia, PA 19118

ABSTRACT

To obtain a molecular basis for the similarities and dissimilarities in the functional, chemical, and biochemical properties between β -casein and the other caseins, a predicted three-dimensional model is presented. The predicted structure was assembled using molecular modeling techniques, as well as secondary structural prediction algorithms, in conjunction with global secondary structural information from Raman spectroscopy. To add validity to this model, the structure was refined using energy minimization techniques to diminish the likelihood of structural overlaps and energetically unfavorable van der Waals contacts arising from the large number of proline residues present in the β -casein sequence. The refined model overall showed a loosely packed, asymmetrical structure with an axial ratio of 2:1. Hydrophobic side chains were uniformly dispersed over one end (C terminal) and the center surface of the structure; the other end (N terminal) was hydrophilic. The hydrophobic section also possessed a large loop through which water could easily travel. Such a suprasurfactant structure could account for the micellar type of hydrophobically driven self-association exhibited by β -casein. Other chemical and biochemical data are in good agreement with the refined structure.

(Key words: casein structure, protein functionality, milk proteins)

Abbreviation key: R_g = radius of gyration, SA = surface area, V = volume.

INTRODUCTION

Previous reports (24, 25) from this laboratory described three-dimensional unrefined structures of κ - and α_{s1} -caseins. These structures were assembled using molecular modeling techniques, sequence-based secondary structural prediction algorithms, and global secondary structural analysis derived from Fourier-transformed infrared and Raman spectroscopy results. The κ -casein model showed two sets of hydrophobic antiparallel β -sheets of nearly equal length, each set containing a proline in a γ -turn as a pivot point for the generation of antiparallel β -sheets. The α_{s1} -casein model showed two domains, one hydrophobic and one hydrophilic, connected by a segment of α -helical structure. The hydrophobic domain of α_{s1} -casein, like that of κ -casein, also contains two sets of hydrophobic antiparallel β -sheets; however, herein one set is approximately twice the size of the other. Such sheet structures were speculated to give some degree of specificity to hydrophobically controlled self-associations, as well as to interactions between κ - and α_{s1} -casein. In addition, the two molecular models showed good agreement with chemical and biochemical information available from the literature (24, 25).

β -Casein represents 36% of bovine casein (14). β -Casein has a variety of functional properties that in some cases are similar to, but mostly differ from, those of α_{s1} - and κ -casein (19). Moreover, the relative amount of β -casein differs significantly among species; e.g., in human milk, in which small amounts of α_{s1} -casein occur, β -casein is the major casein component (19).

The β -casein molecule is a single chain of known sequence (33) with five phosphoserine

TABLE 1. Profile of the β -casein molecule derived from its primary structure.¹

Residues considered	Net charge ²	Charge frequency ³	Average hydrophobicity ³
1 → 43	-16	.55	783
44 → 92	-1.5	.10	1429
93 → 135	+1	.28	1173
136 → 177	+2	.12	1467
178 → 209	0	.13	1738

¹Derived from the data of Ribadeau-Dumas et al. (33).

²Serine phosphate = -2, histidine = +.5; charges (+ or -) per residue.

³Calculated as described by Bigelow (6) and expressed in kilocalories per residue.

residues and a molecular weight of 23,980 Da. The charge frequency, hydrophobicity, and net charge for the various segments of β -casein are presented in Table 1. Analysis of these data indicates that β -casein could be a linear amphiphile and thus much more "soaplike" than α_{s1} -casein. The N-terminal portion of the β -casein molecule (residues 1 → 40) contains the phosphoserine residues and carries essentially all of the protein's net charge as well as most of the protein's potential α -helical residues (13). The C-terminal one-third of the molecule (actually, residues 136 → 209) contains many apolar residues (resulting in its high hydrophobicity) and only two short stretches of potential β -structure as predicted by Creamer et al. (13). β -Casein undergoes an endothermic self-association that reaches a maximum or limiting size depending on the ionic strength (3, 23, 30, 36, 39). The N-terminal concentration of charge, coupled with the highly hydrophobic C terminal, may account for the temperature dependence of this self-association, because hydrophobic interactions are temperature-dependent (18, 19). The self-association of β -casein can be fitted best to a model that describes the association as proceeding through a critical micelle concentration and achieving a limiting size (23, 36, 37, 39). Like α_{s1} -casein, β -casein is insoluble at room temperature (24°C) in the presence of Ca^{2+} at total calcium concentrations below those encountered in milk (18, 19). However, precipitation of β -casein from solution is temperature-dependent, and the calcium- β -caseinate complex is soluble at 1°C at concentrations of up to 400 mM Ca^{2+} (18). Again, this temperature dependence is demonstrated in the release of β -casein from the casein micelle at temperatures between 2 and 4°C (1, 15, 16, 34). In

aqueous solution, β -casein was predicted to occur as a random coil with little or no regular secondary structure (13, 19, 28), which agrees with analysis of the complete amino acid sequence of β -casein, but not with more recent spectroscopic results from this laboratory (7, 8). The proline content of the molecule is high and, on first inspection, appears to be rather evenly distributed (Figure 1A). However, proline residues are concentrated in the C-terminal portion. Recent Raman spectroscopic analysis of the caseins (7, 8) revealed the presence of a high degree of β - or possibly γ -turns in the molecule. In the building of molecular models, for the α_{s1} - and κ -casein models (24, 25), a correlation was made between the proline content and the turn content as measured by spectroscopy. Thus, the investigation of the possible role of proline in turns in β -casein seemed to be of interest.

This paper is intended to construct a possible predicted three-dimensional structure for β -casein, using, as in the studies (24, 25) of κ - and α_{s1} -casein, molecular modeling, sequence-based prediction programs, and Raman spectroscopy results. However, unlike the previous studies (24, 25), this structure is refined via energy minimization techniques, using a Kollman force field (40) to evaluate the effects of sequential proline residues in the primary sequence and to produce a working molecular model for this protein.

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

MOLECULAR MODELING OF β -CASEIN

sequence-based predictive techniques, primarily those of Chou and Fasman (9), Garnier et al. (22), and Cohen et al. (10, 11). Assignments of secondary structure (α -helix, β -sheet, or β -turn) for the amino acid sequences were made either when predicted by more than one method or when 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. (4) and Ananthanarayanan et al. (2).

Molecular Modeling

The three-dimensional structure of β -casein was approximated using molecular modeling methods with an Evans and Sutherland PS390 interactive computer graphics display driven by Sybyl software (Tripos, St. Louis, MO). The construction of this model was accomplished by drawing the predicted behavior of the polypeptide chain from its amino acid sequence and reconciling these predictions with spectroscopic data. In building the macromolecule, we used a library or a dictionary of

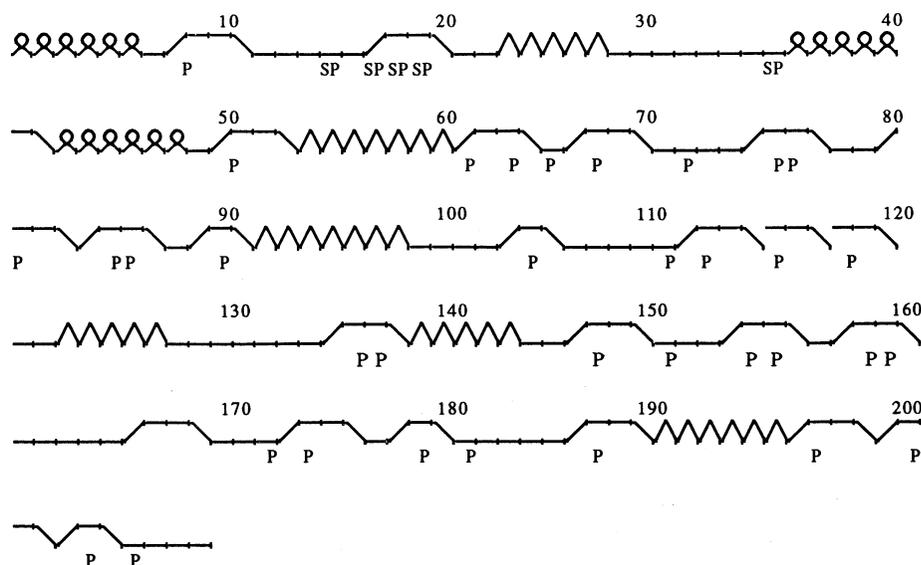
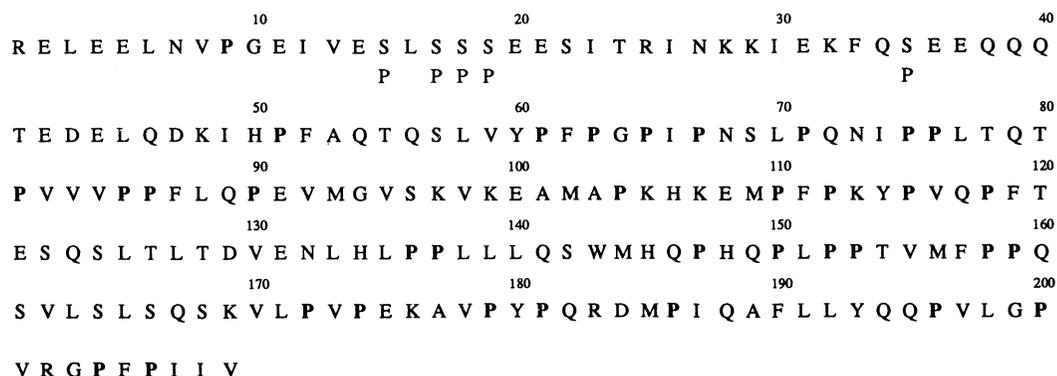


Figure 1. A) Sequence of β -casein A²; proline residues are boldface; capital P below S indicates serine phosphate. B) Summary of initial secondary structural assignments made for β -casein A²; P denotes proline; SP denotes serine phosphate.

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 β -casein A² model followed the procedures previously developed for the models of α_{s1} - and κ -caseins (24, 25). The protein was built in segments, 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 the totally random configuration. The Sybyl subroutine SCAN was used, on the side chains only, to adjust torsional angles and to relieve bad van der Waals contacts. The individual pieces were then joined together to produce the total polypeptide model and readjusted.

Molecular Force Field Energy Minimization

Quantum mechanical calculations have greatly facilitated the development of structure-function relationships for small molecules, i.e., lowest potential energy versus geometry. However, many problems of biological interest, such as protein conformation, still require elementary model empirical energy functions. Although the functions are crude, they have been applied successfully to the study of hydrocarbons, oligonucleotides, peptides, and proteins (40).

Herein, the potential energy model is simply described as a collection of overlapping balls for the atoms with given van der Waals radii connected by springs, 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 that measure bonded and nonbonded interactions quantitatively. 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 a Born-Oppenheimer energy surface for a molecule or system of molecules can be replaced by an analytical function (40). The potential energy function chosen is generally given as a sum of bond energies and nonbonded interaction terms:

$$E_{\text{total}} = \sum_{\text{bonds}} K_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{1}{2}K [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \frac{B_{ij}}{R_{ij}^{12}} - \frac{A_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \quad [1]$$

The first three terms 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 nonbonded van der Waals and electrostatic interactions. Although more complex functions are often used to describe bond stretching and bending or different forms of the nonbonded repulsions, the current version of AMBER (40) by Kollman in Sybyl uses the functional form just presented, plus a 10 to 12 hydrogen bond function if so chosen by the user. 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 Models with Multiple Proline Residues

Various methodologies for sequence-based secondary structural predictions are currently available (9, 10, 11, 22), and these methods have been applied to β -caseins (3, 13). In our initial approaches, a sequence-based prediction was generated from each of three basis sets for

MOLECULAR MODELING OF β -CASEIN

 TABLE 2. Type of proline turn and sequence in β -casein.

β -Turn ¹	γ -Turn ²	Pro-Pro Chain	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.

β -casein A². The results were ambiguous; for some portions of the sequences, structures (such as α -helix) were consistently predicted, whereas, in other portions, they were not. The relatively high proline content of β -casein (Figure 1A) poses a problem because this residue, although occasionally found in both β -sheet and α -helical structures, is generally not favorable to either structure. In attempting to generate a consensus sequence-based prediction, we first focused on solving the proline problem in a way commensurate with known behavior of this residue in proteins and model peptides (2, 4, 35).

In model peptides, proline frequently occupies the second position of either a four residue β -turn or a three residue γ -turn (2, 4, 35). Moreover, recent Raman and Fourier-transformed infrared spectroscopic analyses of the caseins showed that up to 35% of the residues in all caseins appear to be in β - or γ -turns (7, 8). To assess the probability that proline residues in the caseins might be located in such turns, we examined the tetrapeptide sequences containing proline in the second position for similarity with peptides known to

form β - and γ -turns (2, 4). Most of the proline residues were predicted to be in a turn of some type. When the models were built, normal β -turns containing proline residues sometimes resulted in unfavorable van der Waals contacts with surrounding residues, leading us to assign single proline residues initially to either the β - or the γ -turn conformation, depending on the neighboring residues (Table 2).

Many proline residues of β -caseins, unlike those of α_{s1} - and κ -casein, 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. In a previous study of κ -casein (24), 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 used; similar structures for β -casein sequences given in Table 2 were modeled in this work. These β -casein peptide models result in structures that do not unduly constrain the polypeptide chains and that have minimized energies of approximately -8 kcal per residue, which is equivalent to that attained by energy minimization of X-ray crystallographic structures.

TABLE 3. Comparison of adjusted sequence-based predictions and refined structure with Raman data.

Sample		Helix	β -Structure	Turns	Unspecified
		(%)			
β -Casein	Raman ¹	6 - 14	20 - 23	31 - 36	33 - 36
	Initial structure	8	18	40	39
	Final model	10	20	34	34

¹D₂O solution (8).

TABLE 4. Dihedral ϕ and ψ angles assigned to specific conformational states in the initial structures for β -casein.

Structures	Angle assignments	
	ϕ	ψ
α -Helix	-58	-47
β -Sheet	-139	135
β -Turn N	-139	135
N + 1	-60	-30
N + 2	-90	0
N + 3	-139	135
γ -Turn N	135	-69
N + 1	-75	59
N + 2	81	-126
X-Pro-Pro-X		
N	135	69
N + 1	-82	59
N + 2	-82	59
N + 3	81	-126
Pro-X-Pro		
N	-82	59
N + 1	-72	83
N + 2	-82	59

β -Turns, other than those based on proline, were predicted by the methods of Chou and Fasman (9) and Cohen et al. (10, 11). The total number of amino acid residues predicted initially to be in turns (40%) was only slightly in excess of the 30 to 36% predicted by spectroscopic methods (Table 3).

Similarly, "consensus" scores for α -helix and β -sheet were obtained by choosing from those regions having the highest predicted probability of a given structure to yield values in accord with Raman data (8). In this case, all residues previously assigned to proline-based turns were eliminated first from consideration in α -helical or extended β -structures. Finally, all residues not included in these periodic structures were then considered to be in an extended aperiodic conformation. The initial conformational assignments for β -casein are shown with its sequence in Figure 1. The β -casein model resembles that of Creamer et al. (13) except that the β - and γ -turns and proline assignments are included.

Generation of Three-Dimensional Models

The secondary structural assignments (Figure 1B) that had been reconciled with Raman spectroscopic data were used to generate

TABLE 5. Summary of the energy components for energy-minimized β -casein A² model.

Energy	(kcal/mol)
Bond-stretching	43.696
Angle-bending	627.030
Torsional	538.268
Out of plane bending	19.060
1-4 van der Waals	311.553
van der Waals	-1003.735
1-4 Electrostatic	2234.858
Electrostatic	-5454.868
Hydrogen bond	-58.051
Total	-2742.188

the three-dimensional models. Idealized ϕ and ψ angles assigned initially for each structural element are given in Table 4. The protein model was constructed starting with the N-terminal end and proceeding toward the C-terminal end as previously described (24, 25). The protein model was examined visually for "bad contacts" and adjusted for unfavorable van der Waals interactions of the side chains. The initial backbone conformation was maintained throughout this procedure. This procedure is analogous to the mechanism whereby the protein is synthesized in vivo from N \rightarrow C terminal and presumably folds after insertion into the lumen of the endoplasmic reticulum (19).

Energy Minimization of Constructed Model

The total initial structure was energy minimized using the Kollman force field (40) potential. 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 hydrogen atoms. This united atom approach is widely used when dealing with proteins in excess of 30 residues (40). Electrostatic interactions were added to the calculation by using united atom partial charges according to Weiner et al. (40). A nonbonded cutoff of 8 Å was employed for van der Waals and electrostatic nonbonded interactions. The structure was minimized to a limit of ± 1 kcal; the results for each type of energy, as delineated in Equation [1], are presented in Table 5. This total energy cor-

MOLECULAR MODELING OF β -CASEIN

responds to -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 and shown in orthogonal views in Figure 2. The hydrophobic side chains are green; serine phosphates and aspartic and glutamic acid side chains are red; and lysine and arginine groups are purple. The rest of the molecule is cyan except for plasmin and chymosin cleavage sites (20, 38); for these residues, the backbone and side chain atoms are red-orange and orange, respectively. Figure 2 shows that, although the structure of β -casein A² appears to be more compact than the previously derived structures for κ - and α_{s1} -caseins (24, 25), the molecule still contains a good deal of asymmetric character. The backbone of the model demonstrates loops through which water can easily pass. Hence, β -casein cannot be categorized as either a globular or a random coil structure. On inspection, the model can be thought to have a "crablike" appearance with two large, distorted hydrophilic arms. The lower arm (at 4 o'clock position) results from multiple proline-based turns (residues 85 to 119). The other hydrophilic arm (residues 28 to 55) results from a combination of secondary structures that includes extended sheet, turns, and an α -helical segment. Phosphoserine-35 is found in this arm. Both arms contain relatively high charge frequency (Table 1) and well-characterized plasmin cleavage sites that occur at residues 28 to 29, 105 to 106, and 107 to 108 (17).

Figure 2 shows that the overall shape of the molecule is asymmetric and can be approximated by a prolate ellipsoid with an axial ratio of 2:1. Furthermore, one end of the structure is predominately hydrophilic, as just noted, and the rest consists mostly of a highly hydrophobic domain (left side of Figure 2 bottom). 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 invari-

ably nonpolar, as are 60% of the 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 do in 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, unstranded β -sheets (24, 25). This model for β -casein gives the impression of a large, distorted surfactant molecule. A dipole moment of 1557-Debye units and a net charge of -18.25 could be calculated from this structure using the Kollman united atom partial charges (40), whereas 815-Debye units would be calculated from a spherically symmetric equivalent structure with the same net charge. A radius of gyration of 23 Å was calculated by assuming a solid prolate of revolution with an axial ratio of 2:1. A backbone structure with labeled proline residues and a stereoscopic view are also presented in Figure 3, A and B, respectively. The plasmin and chymosin cleavage sites (labeled in Figure 3C) all show accessibility for surface interactions with enzymes. Tryptophan-143 is relatively exposed in the monomer model. Pearce (31) demonstrated a blue shift of fluorescence when β -casein self-associates with increasing temperature. Thus, tryptophan-143 is exposed in the monomer but buried in the polymer, accounting for the blue shift reported to occur on aggregation (31). The phosphoserine residues, which are located in a turn region, as was the case for α_{s1} -casein (25), are also contained in the hydrophilic end of the structure. An extended turn region that resembles a crab's head (Figure 2 and 3) ranges from residue 14 to 22 and contains four of the five phosphoserines in β -casein. These phosphoserines are in a loop, on the surface, and readily exposed for interaction with Ca^{2+} or dephosphorylation by phosphatase (19).

Secondary Structure Analysis

The energy-minimized three-dimensional structure of β -casein A² was compared with reported secondary structural analysis via Raman spectroscopy (Table 3). Figure 4 shows a Ramachandran plot of ϕ and ψ backbone dihedral angles (open circles) calculated from the

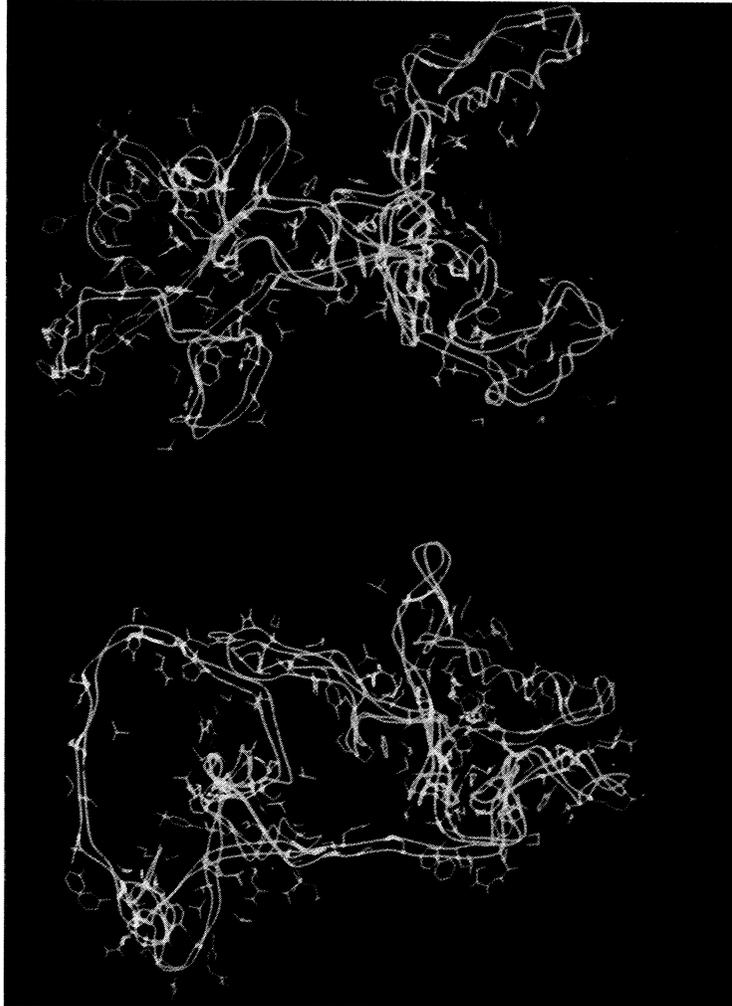


Figure 2. Orthogonal views (90° rotation in y-z plane) of the three-dimensional model of β -casein-A². The hydrophobic side chains are green, acidic side chains are red, and basic side chains are purple. Side chain atoms of plasmin cleavage sites are red-orange, chymosin cleavage sites are orange, and the C-terminal valine 209 is magenta. All other atoms are cyan, and the backbone atoms are replaced by a double ribbon that traces the backbone and is yellow. The bottom view is used in all specific discussions in the text.

MOLECULAR MODELING OF β -CASEIN

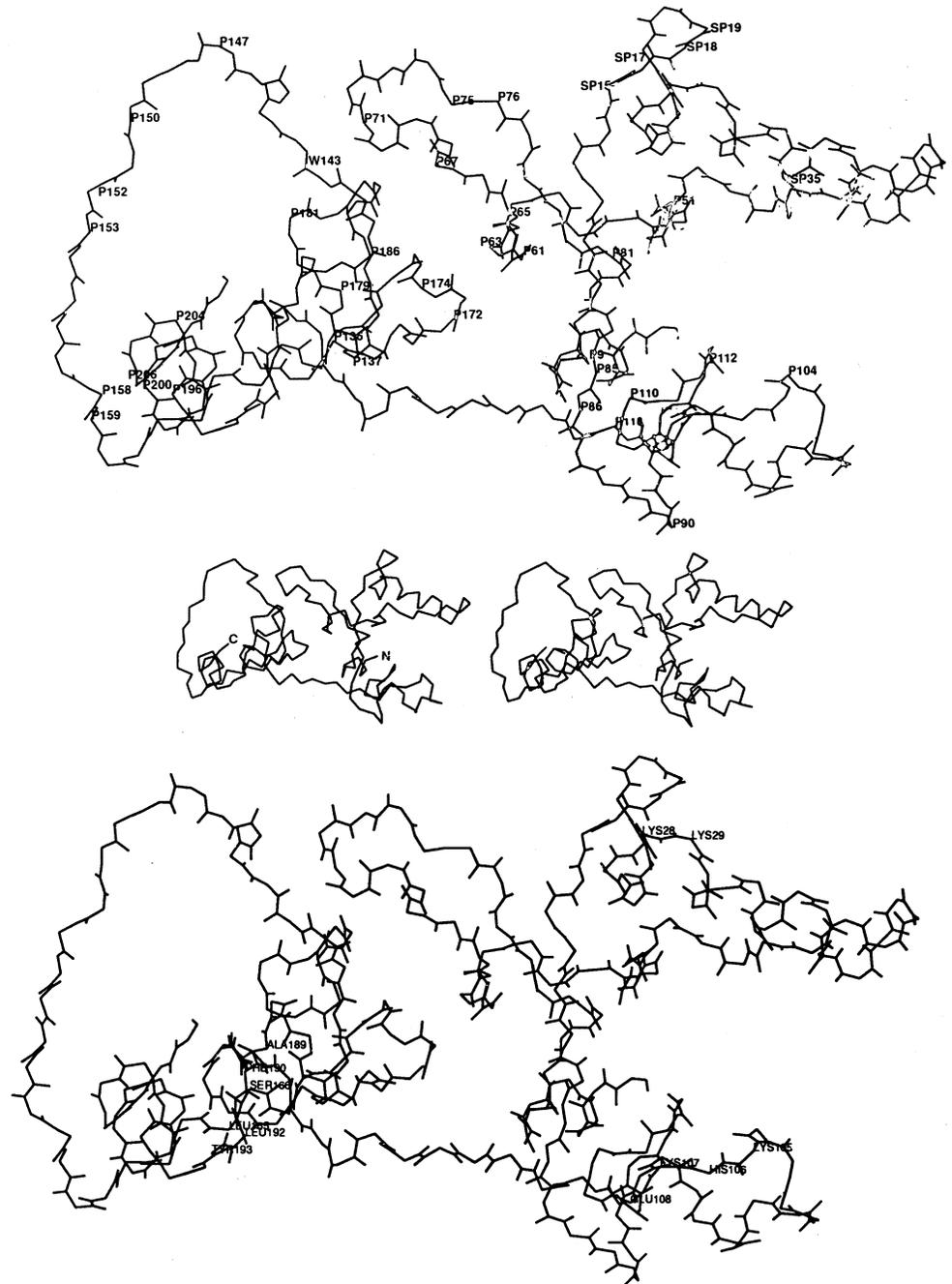


Figure 3. A) Chain trace of β -casein-A² with prolines (P), phosphoserines (SP), and amino (N) and carboxyl (C) terminals indicated. B) Stereoscopic view, relaxed, of β -casein A² structure, showing α -carbon atoms without side chains. C) Same orientation as in A; plasmin and chymosin cleavage sites are shown with residues labeled.

refined β -casein A² structure using the Tripos' Sybyl molecular modeling software; areas of acceptable ϕ and ψ angles for secondary structure are denoted (35). From this plot, the number of residues present within the limits of the particular structure can be calculated to obtain the global amounts of α -helix, β -sheet, or turns. However, this type of analysis does not take into account the required minimum num-

ber of sequential residues to sustain a periodic structure (five to six). In addition, ϕ and ψ angles may not exactly represent the secondary structure because of changes in backbone bond lengths or hydrogen bonding. Hence, visual inspection of the secondary structure by use of a chain trace or ribbon, as in Figure 3A, should also be employed to analyze the secondary structure of a model.

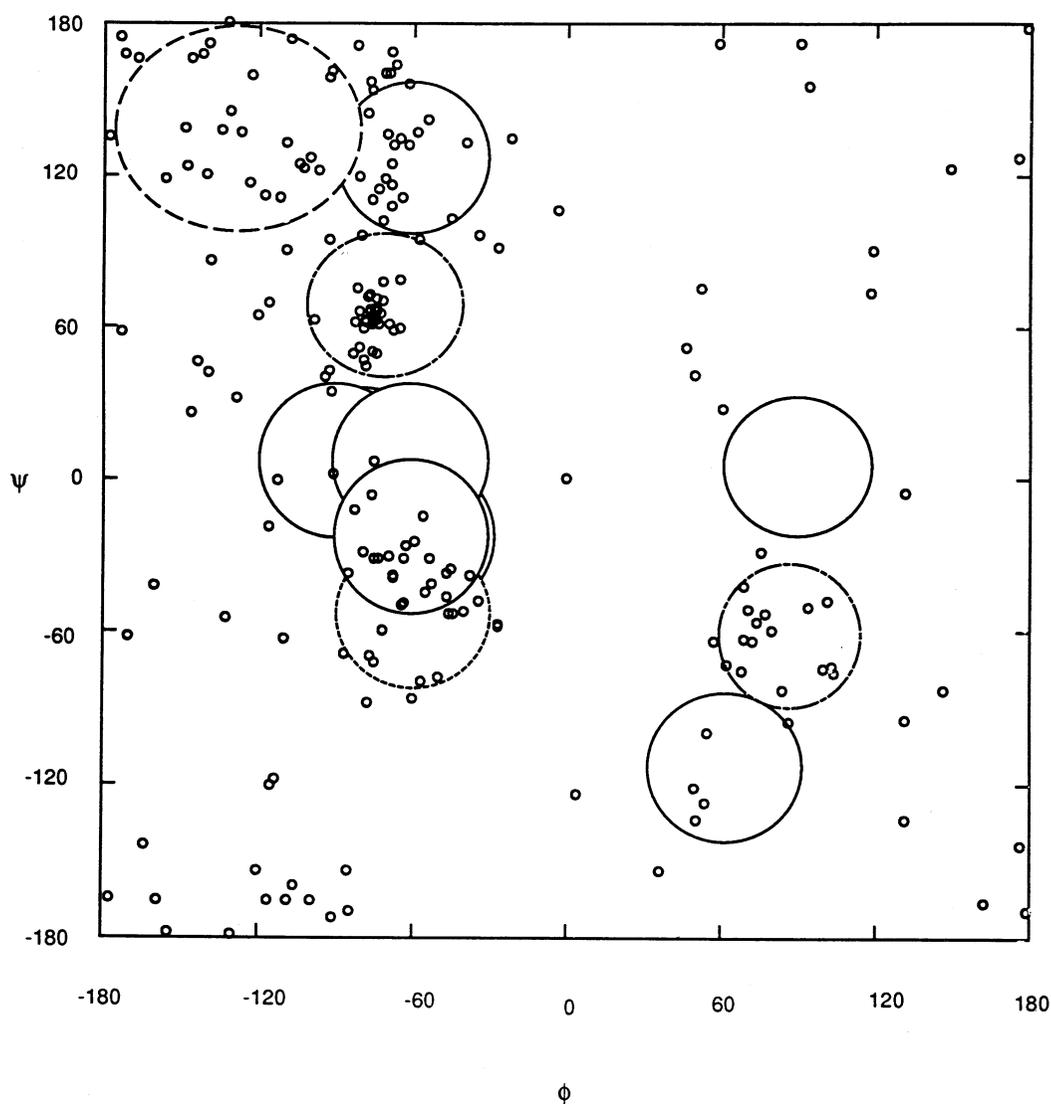


Figure 4. Ramachandran plot of ψ and ϕ angles calculated from energy-minimized structure of β -casein A². Area for β -sheet structure (---), for α -helix (···), for β -turns (—), and for γ -turns (-·-·-).

MOLECULAR MODELING OF β -CASEIN

The procedure described was used to analyze the β -casein A² structure; the global secondary structural results are in good agreement with those obtained via Raman spectroscopy results in D₂O (Table 3). In addition to these classical structures, other types of nonclassical structure are predicted. One type is the β -spiral arising from multiple proline turns, as observed for residues 196 to 206 and 61 to 71 (Figure 3, A and B). Another type of nonclassical structure is the distorted hydrophilic arm, discussed in the previous section, which also results from multiple proline turns contained within residues 85 to 90 and 100 to 120.

Plasmin and Chymosin Cleavage Sites

Several investigators [for reviews, see (17, 20, 38)] determined the sites of β -casein that are cleaved by the proteolytic enzymes, plasmin and chymosin. Further examination of the β -casein A² three-dimensional model is needed to ascertain whether these proteolytic sites are buried or exposed on its surface. The principal plasmin cleavage sites, residues 28 to 29, 105 to 106, and 107 to 108, are red-orange on the model (Figure 2). All appear to be readily exposed on the surface of the proposed structure. The cleavage sites are all located on the hydrophilic side of the molecule and particularly on the long, distorted arms. Hence, all sites not only are solvent-accessible on the surface but also are sufficiently exposed that plasmin in solution would readily hydrolyze these sites as reported (17, 20, 38).

The chymosin cleavage sites are also exposed on the monomer surface but are located on the hydrophobic (left) end of the proposed structure (see Figure 2). This model predicts that hydrophobic interactions, because of temperature-dependent self-association or interactions with other caseins, inhibit chymosin action on β -casein. Visser (38) and Creamer (12) showed that successful chymosin action on β -casein in solution occurs predominately at either low temperatures, at which hydrophobically driven self-association is minimized, or at lower pH. Berry and Creamer (5) and Creamer (12) successfully limited hydrolysis to residues 189 to 190 on β -casein at 2°C. The resulting large polypeptide, i.e., residues 1 to 189, showed a marked decrease in self-association relative to the parent

β -casein by gel-permeation chromatography. In our model, residues 189 to 190 are located in a β -sheet structure (Figures 1 and 2), which is followed by a unique β -spiral (Figure 3, A 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 when hydrophobic self-association occurs except for the C-terminal valine. The relationship of this β -sheet-spiral segment to β -casein self-associations is discussed in more detail in the next section.

Correlation with Solution Physicochemical Studies

Payens (29), Payens and van Markwijk (30), and Schmidt (36) determined the sedimentation coefficient of β -casein monomer at 2°C to be 1.5 S (Svedberg units) by extrapolation to zero protein concentration, which is much lower than that obtained for a globular protein of the same molecular weight, i.e., 2.4 to 2.5 S (24, 25). Kumosinski and Pessen (26, 27) developed a methodology for the prediction of sedimentation coefficient using radius of gyration (R_g), surface area (SA), and volume (V) parameters determined from either small-angle X-ray scattering or from X-ray crystallographic structures. The proposed three-dimensional structure for β -casein A² has an R_g of approximately 23 Å and dimensions of 21 by 42 Å, assuming a prolate ellipsoid of an axial ratio of 2 to 1. Because the Sybyl molecular modeling software does not provide adequate SA and V calculations, a volume of 77,600 Å³ was estimated from the dimensions of the prolate ellipsoid model by assuming an SA:V of .2 Å⁻¹. This latter value was calculated by Kumosinski and Pessen (26, 27) for pepsin, which has a large SA. Using this SA:V ratio in conjunction with the molecular weight and partial specific volume, a sedimentation constant of 1.7 S could be calculated for the β -casein. However, pepsin is a globular protein, and the proposed model of β -casein could have a much higher SA:V ratio because of the apparent rugosity of its SA (Figures 2 and 3). Hence, this calculated 1.7 S is an upper limit, and, in reality, a more precise SA calculation would yield a sedimentation coefficient much closer to the experimentally determined 1.5 S (30, 36).

MOLECULAR MODELING OF β -CASEIN

initiated to form various aggregates. To limit the number of possible structures, only side chain-side chain hydrophobic docking was allowed, which is consistent with the hydrophobic self-association character of the system. In addition, we utilized the results of Berry and Creamer (5) and Creamer (12), which indicate that residues 189 to 190 and the C-terminal valine remain exposed on the aggregate structure and that elimination of the β -sheet-spiral structure of residues 189 to 207, discussed in the previous section, would destabilize these aggregates. Figure 5 shows an energy-minimized tetramer structure with a "ribboned" backbone, which is consistent with the constraints described and yields a calculated R_g of 43 Å, which is in reasonable agreement with the experimental value (3) of 46 Å. Even an end to end dimer model, with the hydropho-

bic ends docked together and the hydrophilic portions at opposite ends, yielded an R_g of only 34 Å. As seen in Figure 5, the tetramer structure is approximately a rod of dimensions 84×42 Å with centrally located hydrophobic groups for further self-association and hydrophilic groups at each end. This resulting structure (Figure 5) allows for easy accessibility of the chymosin cleavage site 189 to 190, as well as destabilization of the tetramer structure, when the 190 to 206 hydrophobic β -sheet-spiral structure is eliminated by the action of carboxypeptidase or chymosin (5, 12).

Tai and Kegeles (37) showed via sedimentation experiments that β -casein aggregation can be described by a model based upon "micelle" formation at room temperature (24°C), with a degree of self-association of 20. Such a 20-mer aggregate can be constructed



Figure 5. Proposed energy-minimized tetramer model of β -casein A_2 with radius of gyration of 43.1 Å. The hydrophobic side chains are green, acidic side chains are red, and basic side chains are purple. Side chain atoms of plasmin cleavage sites are red-orange, chymosin cleavage sites are orange, and the C-terminal valine-209 is magenta. All other atoms are cyan, and the backbone atoms are replaced by triple ribbons that trace the backbones and are yellow, magenta, orange, and red.

from the tetramer model of Figure 5, by parallel placement of four more tetramers structures around a centrally located structure. The resulting structure would allow for only hydrophobic side chain contacts and would have a dimension of approximately 84×84 Å. However, this structure could not be approximated by a sphere or rectangle model because of the diagonal cross-section dimension of only 21 Å; four more tetramers are needed to approximate a smooth surface model. Therefore, an R_g of 73 Å, as determined by Andrews et al. (3), would be plausible using the 20-mer model.

A unique feature of the tetramer model is that further aggregation caused by hydrophobic side chain contacts would easily be initiated when salt is bound to the hydrophilic end portion of the structure, thereby minimizing electrostatic repulsion. Such ionic strength dependence of increased aggregation has been noted by many investigators for β -casein (23, 37, 39).

CONCLUSIONS

We have presented an energy-minimized predicted three-dimensional structure of β -casein A^2 using a strategy based on secondary structure from sequences, global secondary structural results from Raman spectroscopy, and molecular modeling techniques. This structure is in good agreement with biochemical cleavage results for plasmin and chymosin action on β -casein but in less agreement with experimentally derived results from sedimentation and small-angle X-ray scattering experiments; however, these latter data are incomplete. The model also provides a molecular basis for the temperature-dependent self-association of β -casein. Because the model is static and not dynamic, it does not allow for a choice between the proposed mechanisms of conformational changes between 2 and 8°C. The monomer model is perhaps a more structured molecule than that existing below 2°C, and therefore the polymer model does not depend on whether or not conformational changes precede or are the result of aggregation. However, this structure must be viewed as a working model with the flexibility to be changed as more precise experiments are performed to ascertain the validity and predictability of this three-dimensional structure. In

future work, molecular dynamics calculations will be performed on this structure to test its stability when a kinetic energy equivalent to a bulk temperature is applied. The current model is a static in vacuo structure, and only a dynamic structure, not a static structure, can be directly compared with experimentally derived solution results, especially in the case of the conformationally related aggregation of β -casein between 2 and 8°C. However, the current model represents a starting point and is assumed to represent an average dynamic structure. This structure serves to allow comparisons with physical chemical data and information from small-angle X-ray scattering.

REFERENCES

- 1 Ali, A. E., A. T. Andrews, and G. C. Cheeseman. 1980. Influence of storage of milk on casein distribution between micellar and soluble phases. *J. Dairy Res.* 47:371.
- 2 Ananthanarayanan, V. S., S. K. Brahmachari, and N. Pattabiraman. 1984. Proline-containing β -turns in peptides and proteins: analysis of structural data on globular proteins. *Arch. Biochem. Biophys.* 232:482.
- 3 Andrews, A. L., D. Atkinson, M.T.A. Evans, E. G. Finer, J. P. Green, M. C. Phillips, and R. N. Robertson. 1979. The conformation and aggregation of bovine β -casein A. I. Molecular aspects of thermal aggregation. *Biopolymers* 18:1105.
- 4 Benedetti, E., A. Bavoso, B. D. Blasio, V. Pavone, C. Pedone, C. Toniolo, and G. M. Bonora. 1983. Solid-state geometry and conformation of linear, diastereoisometric oligoprolines. *Biopolymers* 22:305.
- 5 Berry, G. P., and L. K. Creamer. 1975. Association of bovine β -casein. Importance of the C-terminal region. *Biochemistry* 14:187.
- 6 Bigelow, C. C. 1967. On the average hydrophobicity of proteins and the relation between it and protein structure. *J. Theor. Biol.* 16:187.
- 7 Byler, D. M., and H. M. Farrell, Jr. 1989. Infrared spectroscopic evidence for calcium ion interaction with carboxylate groups of casein. *J. Dairy Sci.* 72: 1719.
- 8 Byler, D. M., H. M. Farrell, Jr., and H. Susi. 1988. Raman spectroscopic study of casein structure. *J. Dairy Sci.* 71:2622.
- 9 Chou, P. Y., and G. D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* 41:45.
- 10 Cohen, F. E., R. M. Abarbanel, I. D. Kuntz, and R. J. Fletterick. 1983. Secondary structure assignment for α/β proteins by combinatorial approach. *Biochemistry* 22:4894.
- 11 Cohen, F. E., R. M. Abarbanel, I. D. Kuntz, and R. J. Fletterick. 1986. Turn prediction in proteins using a pattern-matching approach. *Biochemistry* 25:266.
- 12 Creamer, L. K. 1976. A study of the action of rennin on β -casein. *N.Z. J. Dairy Sci. Technol.* 11:30.

MOLECULAR MODELING OF β -CASEIN

- 13 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.
- 14 Davies, D. T., and A.J.R. Law. 1980. Content and composition of protein in creamery milks in South-West Scotland. *J. Dairy Res.* 47:83.
- 15 Davies, D. T., and A.J.R. Law. 1983. Variation of the protein composition of bovine casein micelles and serum caseins in relation to micellar size and temperature. *J. Dairy Res.* 50:67.
- 16 Downey, W. K., and R. F. Murphy. 1970. The temperature dependent dissociation of β -casein from bovine casein micelles. *J. Dairy Res.* 37:361.
- 17 Eigel, W. N., J. E. Butler, C. A. Ernstrom, H. M. Farrell, Jr., V. R. Harwalker, R. Jenness, and R. M. Whitney. 1984. Nomenclature of the proteins of cows' milk: 5th revision. *J. Dairy Sci.* 67:1599.
- 18 Farrell, H. M., Jr., T. F. Kumosinski, P. Paulaski, and M. P. Thompson. 1988. Calcium induced association of the caseins: a thermodynamic linkage approach to precipitation and resolubilization. *Arch. Biochem. Biophys.* 265:146.
- 19 Farrell, H. M., Jr. and M. P. Thompson. 1988. The caseins of milk as calcium binding proteins. Page 117 in *Calcium Binding Proteins*. M. P. Thompson, ed. CRC Press, Boca Raton, FL.
- 20 Fox, P. F. 1981. Proteinases in dairy technology. *Neth. Milk Dairy J.* 35:233.
- 21 Garnier, J. 1966. Conformation of β -casein in solution. Analysis of a thermal transition between 5° and 40°C. *J. Mol. Biol.* 19:586.
- 22 Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97.
- 23 Javor, G. T., S. M. Sood, P. Chang, and C. W. Slattery. 1991. Interactions of triply phosphorylated human β -casein: fluorescence spectroscopy and light-scattering studies of conformation and self-association. *Arch. Biochem. Biophys.* 289:39.
- 24 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.
- 25 Kumosinski, T. F., E. M. Brown, and H. M. Farrell, Jr. 1991. Three-dimensional molecular modeling of bovine caseins: α_{s1} -casein. *J. Dairy Sci.* 74:2889.
- 26 Kumosinski, T. F., and H. Pessen. 1982. Calculation of sedimentation constant from SAXS. *Arch. Biochem. Biophys.* 219:89.
- 27 Kumosinski, T. F., and H. Pessen. 1985. Calculation of hydrodynamic parameters from X-ray diffraction and small-angle X-ray scattering data. *Methods Enzymol.* 117:219.
- 28 Noelken, M., and M. Reibstein. 1968. The conformation of β -casein B. *Arch. Biochem. Biophys.* 123:397.
- 29 Payens, T.A.J. 1979. Casein micelles: the colloidal-chemical approach. *J. Dairy Res.* 46:291.
- 30 Payens, T.A.J., and B. W. van Markwijk. 1963. Some features of the association of β -casein. *Biochim. Biophys. Acta* 71:517.
- 31 Pearce, K. N. 1975. A fluorescence study of the temperature dependent polymerization of bovine β -casein A¹. *Eur. J. Biochem.* 58:23.
- 32 Pessen, H., T. F. Kumosinski, H. M. Farrell, Jr., and H. Brumberger. 1991. Tertiary and quaternary structural differences between two genetic variants of bovine casein by small-angle X-ray scattering. *Arch. Biochem. Biophys.* 284:133.
- 33 Ribadeau-Dumas, B., G. Brignon, F. Grosclaude, and J. C. Mercier. 1972. Primary structure of bovine β -casein. Complete sequence. *Eur. J. Biochem.* 25:505.
- 34 Rose, D. 1968. Relation between micellar and serum caseins in bovine milk. *J. Dairy Sci.* 51:1897.
- 35 Rose, G. D., L. M. Gierasch, and J. A. Smith. 1985. Turns in peptides and proteins. *Adv. Prot. Chem.* 37:1.
- 36 Schmidt, D. G. 1982. Associations of caseins and casein micelle structure. Page 61 in *Developments in Dairy Chemistry-1*. P. F. Fox, ed. Appl. Sci. Publ. Ltd., Essex, Engl.
- 37 Tai, M.-S., and G. Kegeles. 1984. A micelle model for the sedimentation behavior of bovine β -casein. *Biophys. Chem.* 20:81.
- 38 Visser, S. 1981. Proteolytic enzymes and their actions on milk proteins. A review. *Neth. Milk Dairy J.* 35:65.
- 39 Waugh, D. F., L. K. Creamer, C. W. Slattery, and G. W. Dresdner. 1970. Core polymers of casein micelles. *Biochemistry* 9:786.
- 40 Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. Force field calculations in computational chemistry. *J. Comput. Chem.* 7:230.