

Comparison of the Three-Dimensional Molecular Models of Bovine Submicellar Caseins with Small-Angle X-Ray Scattering. Influence of Protein Hydration

Received January 12, 1994

To test the applicability of two energy-minimized, three-dimensional structures of the bovine casein submicelle, theoretical small-angle X-ray scattering curves in the presence and absence of water were compared to experimental data. The published method simulates molecular dynamics of proteins in solution by employing adjustable Debye-Waller temperature factors (*B* factors) for the protein, for the solvent, and for protein-bound water. The programs were first tested upon bovine pancreatic trypsin inhibitor beginning with its known X-ray crystal structure. To approximate the degree of protein hydration previously determined by NMR relaxation experiments (0.01 g water/g protein), 120 water molecules were docked into the large void of the κ -casein portion of the structure for both the symmetric and asymmetric casein submicelle models. To approximate hydrodynamic hydration (0.244 g water/g protein), 2703 water molecules were added to each of the above structures using the "droplet" algorithm in the Sybyl molecular modeling package. All structures were then energy-minimized and their solvation energies calculated. Theoretical small-angle X-ray scattering curves were calculated for all unhydrated and hydrated structures and compared with experimentally determined scattering profiles for submicellar casein. Best results were achieved with the 120-bound-water structure for both the symmetric and asymmetric submicelle models. Comparison of results for the protein submicelle models with those for the theoretical and literature values of bovine pancreatic trypsin inhibitor demonstrates the applicability of the methodology.

KEY WORDS: Casein structure; protein functionality; hydration; small-angle scattering.

1. INTRODUCTION

In a recent report from this laboratory, two possible predicted models of a three-dimensional structure of submicellar casein were presented (Kumosinski *et al.*, 1994). These models were built from sequence-based secondary structure predic-

tion algorithms used in conjunction with global secondary structure results determined from Raman and Fourier transform infrared deconvolution spectroscopy experiments. The models were termed asymmetric and symmetric based upon the symmetry of the placement of β -casein into a framework of α_{s1} - and κ -casein. Both the asymmetric and symmetric structures were refined via energy minimization techniques and showed agreement with chemical and biochemical experimental information (Kumosinski *et al.*, 1994). While the symmetric micelle structure yielded a significantly lower energy of formation than the

¹ USDA, ARS, Eastern Regional Research Center, Philadelphia, Pennsylvania 19118.

² To whom correspondence should be addressed.

Reference to a brand or firm name does not constitute endorsement by the U.S. Department of Agriculture over others of a similar nature not mentioned.

asymmetric model as determined by a Kollman force field, no experimental evidence or theoretical hydrophathy energy calculation could be found to favor either structure.

Much information in the literature exists concerning the influence of water and salt on the structure and functionality of casein (Farrell, 1988; Holt, 1992; Schmidt, 1982). However, the amount of water bound to casein, i.e., the hydration value, is still in question. It appears that the hydration value is highly dependent on the experimental methodology used to study water interactions with casein (Farrell *et al.*, 1989, 1990; Holt, 1992). Hydrodynamics and small-angle X-ray scattering (SAXS) results yield extremely high hydration values of 3–6 g water/g protein for casein (Kumosinski *et al.*, 1983). Hydration values for monomeric globular proteins are approximately 0.23 g water/g protein by these same techniques (Pessen *et al.*, 1988). Deuterium nuclear magnetic resonance (DNMR) relaxation experiments yielded much lower hydration values of 0.007–0.01 g water/g protein for caseins, depending on the conditions of the experiments (Farrell *et al.*, 1989). These values are the same order of magnitude as those determined for globular proteins by DNMR techniques (Farrell *et al.*, 1989). Speculation regarding these discrepancies has led to the proposal that hydrodynamics and SAXS experiments measure to varying extents the density of totally hydrated proteins (Farrell *et al.*, 1989, 1990). DNMR relaxation determines, on the other hand, only the strongly bound water since it measures the actual dynamics of water in the presence of a protein (Farrell *et al.*, 1989). Thus, as is the case with globular proteins (Edsall and McKenzie, 1983), several classes of water–protein interactions must exist for the more amorphous caseins.

Previously, only a qualitative comparison of SAXS results with the two refined submicellar structures was given. In particular, the geometric parameters, low electron density values, were successfully contrasted with those calculated from these two predicted structures (Kumosinski *et al.*, 1994). However, the development of a mathematical formalism for rapidly calculating theoretical SAXS profiles from three-dimensional structures by Lattman (1989) makes these quantitative comparisons now possible. The methodology optimizes the fit between the theoretical curve and experimental data by varying the isotropic temperature factor (B) not only for the protein atoms, but

also for the bound water and solvent water atoms, so that some insight into protein–water interactions may also be achieved, without necessitating a lengthy molecular dynamics calculation.

It is the purpose of this study to test the symmetric and asymmetric submicelle structures by quantitatively comparing theoretical and experimental SAXS using the Lattman (1989) program with varying amounts of water stimulating two classes of bound water. The reliability of the program will first be tested with comparably produced theoretical SAXS data from bovine pancreatic trypsin inhibitor (BPTI), which has a known structure. All fits for unhydrated and hydrated casein structures, as well as their corresponding B factors, will be presented and contrasted with the known BPTI structure for determination of the properties of the postulated casein submicelle structure.

2. METHODS

2.1. Conservation of Hydrated Structure

Low-hydrated structures of the refined, energy-minimized casein submicelle models were constructed using a docking procedure on an Evans and Sutherland (St. Louis, MO) PS390 interactive computer graphics display driven by the Tripod Sybyl (Tripod, St. Louis, MO) molecular modeling software on a Silicon Graphics (Mountain View, CA) W-4D35 processor. The docking procedure allowed for individually manipulating the orientation of 120 energy-minimized water molecules in up to four molecular display areas relative to one another. The desired orientation could then be frozen in space and merged into one molecular display area for energy minimization calculation using a molecular force field. The criterion for acceptance of reasonably hydrated structures was determined by a combination of experimentally determined information, i.e., DNMR relaxation results (Farrell *et al.*, 1989, 1990), in combination with the calculation of the lowest energy for that structure. All water molecules with unacceptable van der Waals interactions were eliminated.

For hydrated structures with large amounts of water, the Tripod “Droplet” algorithm was employed. This procedure creates a monomolecular layer of water around an entire structure in an objective manner. In these calculations, a structure with a low hydration value (120 water molecules)

was created using the above docking procedure, then the high-hydration model was generated using the “Droplet” algorithm. Thus, a total of 2723 water molecules could be objectively added to the low-hydrated structure, yielding a total hydration value of 0.244 g water/g protein.

2.2. Energy Minimization. Molecular Force Field

The concept, equation for, and a full description of a molecular force field was given in a previous communication (Kumosinski *et al.*, 1993). In these calculations, a Kollman force field was employed. This force field used electrostatic interactions calculated from partial charges given by Kollman. A united atom approach with only essential hydrogens for reasonable calculation time on the computer was also employed. A cutoff value of 8 Å was employed for all nonbonded interactions. The conjugate gradient technique was employed as a minimization algorithm for all structures in this study.

2.3. Calculation of SAXS Profiles

All small-angle X-ray scattering profiles were calculated for the unhydrated and hydrated structures using a computer program based on an algorithm developed by Lattman (1989). This methodology not only allows for rapid calculation of SAXS profiles, i.e., at least ten times faster than other procedures, but also allows for optimization of the residual between calculated and experimental SAXS profiles using adjustable temperature factors for protein, bound water, and solvent water. The effects of solvent have been modeled by subtracting from each protein atom a properly weighted solvent water molecule. Protein hydrogen atoms are implicitly accounted for using the strategy of Gelin and Karplus (1979).

The scattering profile is given by

$$I(R) = \frac{1}{2} \sum_{n=0}^x \sum_{m=0}^n \varepsilon_m N_{m,n} |G_{m,n}(R)|^2 \quad (1)$$

where

$$G_{m,n}(R) = \sum_j F_j Y_{m,n}(\theta_j, \phi_j) j_n(2\pi r_j R) \quad (2)$$

I is the scattering intensity, and $R = 2 \sin \theta / \lambda$, where θ is the scattering angle and λ is the wavelength of incident radiation. N is the number

of atoms and ε is a constant related to the order (n or m) of the Legendre polynomial used. $Y_{m,n}$ are complex spherical harmonics, j_n are the spherical Bessel functions. The index j runs over all atoms, i.e., protein, protein-bound water, and solvent water, and r , θ , and ϕ are their corresponding spherical atomic coordinates. The expanded structure factor F_j is given by

$$F_j = (\alpha_P f_P + \alpha_B f_B - \alpha_W f_W) \exp(2\pi i r_j R) \quad (3)$$

where α_P , α_{BW} and α_W are the occupancies of the protein atoms, bound water, and solvent water, respectively. The temperature factors, B are related to the structure factors by

$$f_P = f_P^0 e^{-B_P} \quad (4a)$$

$$f_B = f_B^0 e^{-B_B} \quad (4b)$$

$$f_W = f_W^0 e^{-B_W} \quad (4c)$$

where the f^0 are the structure factors in the absence of thermally induced vibrational motion and the B factor compensates for temperature-induced changes (a Debye–Waller constant). The subscripts of P, B, and W represent the atoms due to the protein, bound water, and solvent water, respectively. The units of B are Å².

All calculations using the Lattman program were performed on a VAX 8350 (Digital Equipment, Waterbury, MA) computer. All BPTI calculations took 20 min to complete, whereas submicelle structures required at least 22 hr.

3. RESULTS AND DISCUSSION

3.1. Bovine Pancreatic Trypsin Inhibitor

To test the programs developed for comparison of energy-minimized structures with SAXS profiles (for hydrated and unhydrated casein submicellar structures) and to more fully understand the parameters calculated from the Lattman program, it was important to first use the procedure on a protein molecule with a known hydrated three-dimensional X-ray crystallographic structure. We followed the lead of Lattman (1989) and used the X-ray and neutron crystallographic structure of BPTI, i.e., the 5BPTI file in the Brookhaven protein data bank. The structure obtained from the protein data bank consists of all protein hydrogens as well as 63 water molecules with hydrogens and one bound anion. To mimic the Lattman calculation on casein models, all hydrogens were

removed and the anion was eliminated as well as the three waters associated with this anion. This unrefined structure was presented to the Lattman program using the scattering data provided in the program files, which were originally determined by Pickover and Engelman (1982). To be consistent with the calculation for submicellar casein, only 20 equally spaced data points were used. The resulting calculated B values for the protein B_P , the bound water B_B , and the solvent water B_W , as well as the residual (the variance of the calculated data from experimental data) as a measure of the goodness of fit, are presented in the first row of Table I. The overall profile of the calculated and experimental SAXS results, as filled triangles with a connecting line, are shown in Fig. 1A. The fit of the structure to the SAXS data is acceptable; even the residual value of 0.220 (Table I) represents an error of only 3%. The B values for the bound water and solvent are in reasonable agreement with those calculated by Lattman (1989), i.e., 59 and 72 \AA^2 , respectively; however, the B_P of 125 \AA^2 is much lower than that presented in his paper, i.e., 284 \AA^2 . Presumably, this slight discrepancy may be caused by either the use of a lower number of experimental data points or the elimination of the anion.

Since the casein submicelle models are energy-minimized, and initially contain no water, the effect of refinement of protein structures via energy minimization on the goodness of fit with experimental SAXS profiles was tested using the Lattman procedure. The BPTI structure with 60 waters from X-ray crystallography was energy-minimized using a Kollman force field with

Table I. Temperature Factors for Bovine Pancreatic Trypsin Inhibitor (BPTI) from Small-Angle X-Ray Scattering^a

Refined	Water molecules	B_P (\AA^2)	B_W (\AA^2)	B_B (\AA^2)	R
No	60	125	55	89	0.220
Yes	60	224	58	200	0.160
No	0	-48	-45	—	0.186
Yes	0	-18	60	—	0.133
No	4	-12	-47	102	0.0833
Yes	4	-3	-41	122	0.0731
Yes	202	137	120	35	0.148

^a B_P , Temperature factor for protein. B_W , Temperature factor for free water. B_B , Temperature factor for bound water. R , Residual: deviation of calculated from experimental SAXS data.

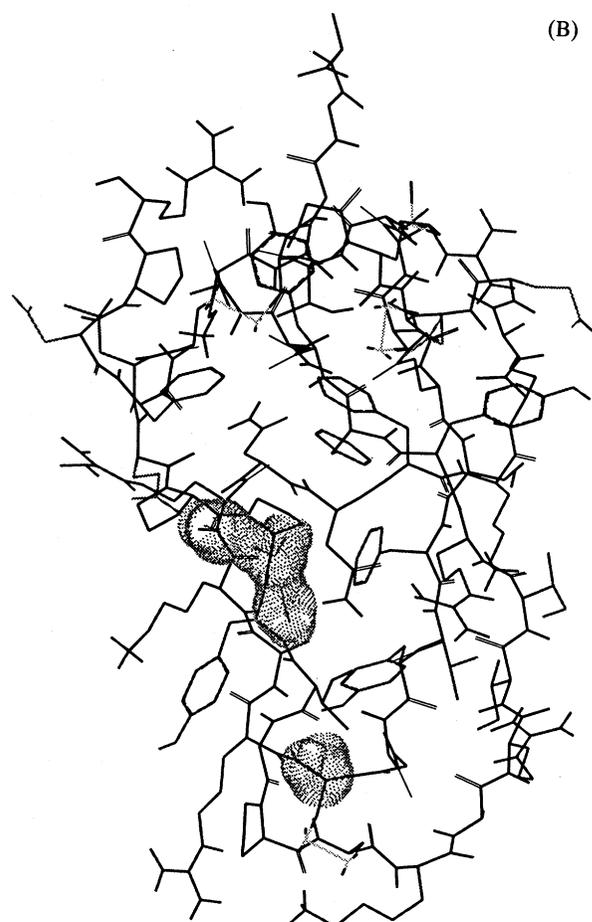
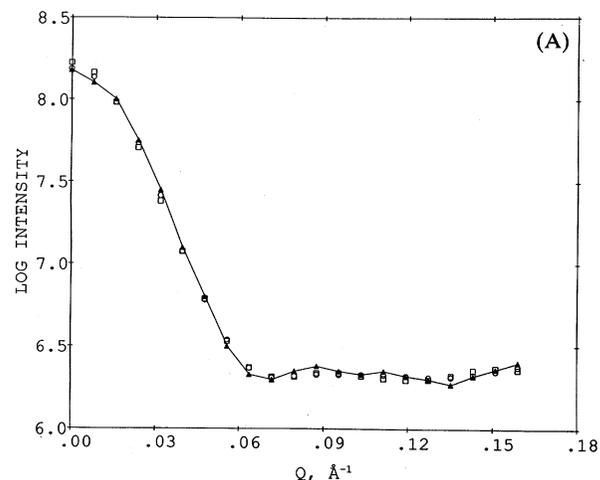


Fig. 1. (A) SAXS profiles of bovine pancreatic trypsin inhibitor (BPTI): filled triangle, with connecting lines, experimental data (15); squares, optimized theoretical curve from X-ray crystallographic structure with 60 waters determined via neutron diffraction; circles, optimized theoretical curve from energy-minimized structure with four internal waters. (B) Structure of BPTI with four internal waters. The structure shows both backbone and all side chains, as well as van der Waals dot surface (gray) of four internal waters.

essential hydrogen added. The resulting energy of -1712.8 kcal, shown in column 3 of Table II, is well below the energy criterion of -10 kcal per residue or water molecule, i.e., -118 kcal/mol, that we have chosen to impose as an acceptable criterion for improvement in the energy value. The resulting refined structure was then presented without hydrogen to the Lattman program and the resulting B and R values are given in row 2 of Table I. The residual (R value) of 0.160 for the refined structure is somewhat lower than that of the unrefined model, while their corresponding B_P and B_B values are much higher, i.e., $B_P = 224$ versus 125 and $B_B = 200$ versus 89 for the refined and unrefined structures, respectively. For $B_P = 224$ there is better agreement with the value of 284 found by Lattman. The reasons for the better fit of the refined over the unrefined structures with the experimental SAXS profiles are not clear at this time. However, since the unrefined structure was determined in the crystal state and the SAXS profiles are obtained in solution where dynamic processes occur, a hypothesis concerning the necessity of a protein adapting to a lower possible energy structure in solution may be offered.

This hypothesis can be further tested at this time by using, via the Lattman program, several unhydrated and hydrated forms of BPTI in their unrefined and refined structures. Here, the refined and original structures of BPTI with no water molecules as well as four internal water molecules [for classification of water molecules associated

with proteins see Edsall and McKenzie (1983)] are used to calculate theoretical SAXS profiles, and the resulting B and R values are given in Table I. The calculated energies for the refined structures of BPTI containing no water molecules and four internal water molecules are shown in Table II. The results of the Lattman program for comparison of the initial and refined structures with SAXS are given in Table I. All temperature factors for the protein B_P are negative as well as all B_W with the exception of the refined structure with no water molecules. Here, no physical interpretation regarding the vibrational motion of atoms of the protein can be involved due to these negative values.

It should be noted at this time that the units of the temperature factors are \AA^2 and the square root of these parameters usually reflects the root mean square displacement of vibrating atoms from their center of mass due to thermal energy. However, the B_P values all become less negative as each structure is refined due to energy minimization. Notably the corresponding values of R decrease when energy minimization is applied to a given complex of BPTI and water. Hence, it appears that energy minimization of structures always improves the goodness of fit between theoretical and experimentally determined SAXS profiles when using the Lattman methodology. Such results indeed added confidence to our present studies of using refined submicelle structures at various hydration levels for comparison of theoretical and experimental SAXS profiles.

Finally, the effect of the variation of hydration levels on R values can be assessed for this methodology. We will now utilize only the refined structures of the hydrated BPTI with various amounts of bound water. The final energies after energy minimization of BPTI at various hydration levels are presented in Table II. Here the total energies of the unhydrated structure, the structure with 60 water molecules from X-ray crystallography, the structure with only four internal water molecules again from the X-ray structure, and the latter structure with the addition of 198 water molecules using the Sybyl droplet algorithm are compared. As can be seen in Table II, the total potential energy decreases with increasing amounts of water in all structures except for that containing the four internal water molecules. The reason for this discrepancy may lie in the fact that these four water molecules occur within the internal structure

Table II. Energy for Hydrated Bovine Pancreatic Trypsin Inhibitor Complex

	Energy for given number of water molecules (kcal/mol)			
	0	60	4	202
Bond stretching	6.4	38.3	10.0	60.8
Angle bending	47.7	60.3	51.7	76.9
Torsional	94.5	94.5	93.3	103.3
Improper torsional	7.0	6.2	7.1	6.9
1-4 Van der Waals	98.1	96.5	96.9	94.8
Van der Waals	-384.8	-177.5	-328.6	-13.0
1-4 Electrostatic	506.4	503.4	501.9	498.6
Electrostatic	-1191.3	-2303.5	-1243.9	-3087.4
H bond	-15.5	-31.1	-20.1	-30.7
Total energy	-831.6	-1712.8	-831.7	-2289.7

of BPTI, and are hydrogen-bonded (Edsall and McKenzie, 1983). In Fig. 1B the structure of BPTI is presented as a "wire frame" model, and the four internally bound water molecules are represented by dot van der Waals surfaces. It is almost surprising to see in Fig. 1B that the four water molecules are trapped within an internal area of the protein which is predominantly hydrophobic by inspection (data not shown) and not hydrophilic as would be normally expected for protein-water interactions. Comparison of the unhydrated structural energy, column 2, and the four-bound-water molecule structural energy, column 3, of Table II, shows that the two structures have approximately the same total energy of about -831 kcal/mol even though the hydrated structure has a lower hydrogen bonding energy of around -20 kcal/mol and an electrostatic energy of about -53 kcal/mol lower than its corresponding energy-minimized unhydrated form. However, this decrease in -53 kcal/mole is counteracted by an increase in van der Waals energy of about 60 kcal/mol. The stabilization in electrostatic and hydrogen bonding energy accompanied by a destabilizing van der Waals energy for the addition of four water molecules could be consistent with the water internally located near hydrophobic side chains and hydrogen-bonded to the protein backbone structure. Such water molecules could be rigidly bound to the protein and would not possess fast local motions as do waters bound at the surface of the protein. This is in agreement with previous analysis of BPTI (Edsall and McKenzie, 1983).

The addition of more water molecules (56 or 198) on the protein surface appears to have an overall stabilization in the total energy of the hydrated protein (Table II). In fact, comparison of these energetics of the 60- and 202-water structures (Table II) corroborates this assumption since no van der Waals energy loss is observed over the unhydrated form. It should be noted that the 60-water-molecule structure contains the four internal as well as 56 surface water molecules as determined via neutron diffraction results. The highly hydrated model (202 total water molecules) was constructed from the model of four internally bound water molecules followed by the use of the droplet algorithm, which added a monomolecular layer of surface water, i.e., 198 water molecules + 4 = 202. Using the total energy of the BPTI-four-bound-water molecule complex as a baseline,

hydration energies of -881.06 and -1458.0 kcal/mol of BPTI can be calculated for the 60- and for the 202-water molecule complexes, respectively. However, the large hydration energy for the 202-water-molecule model represents only a stabilization of -7.3 kcal/added water molecule, while the 60-water-protein model yields a more stabilizing value of -14.5 kcal/added water molecule. Hence, the addition of excess water molecules to a protein surface (especially when a random process is employed as in the droplet algorithm) does not necessarily add to the stabilization of the protein structure. Presentation of these refined hydrated structures to the Lattman procedure for comparison with experimentally determined SAXS profiles would be of interest in this vein.

The results of the optimization of experimental and theoretical SAXS profiles for the refined unhydrated and various hydrated structures of BPTI are given in Table I. The lowest value of the residual R between the experimental and theoretical SAXS profiles is obtained for the refined structure of BPTI with four bound internal water molecules. The fit of the theoretical SAXS profile for the structure is presented in Fig. 1A as open circles. Comparison of the 60-water-molecule-BPTI structure, depicted as squares in Fig. 1A, and of the four-water-molecule structure (open circles) with the experimental data (filled triangles) shows only slight differences at high Q values, but the deviations are mostly at the very low Q values where the radius of gyration and molecular weight are calculated.

Overall, even the 60-water-molecule structure shows a satisfactory fit between theoretical and experimental SAXS profiles even though its R value is two times larger than the refined internally bound four-water-molecule-BPTI structure (Table I). However, since the B value of the protein is negative for the four-internally-bound water molecule-BPTI structure, we would conclude that some more bound waters are necessary for proper simulation of the data. Using the 202-water-molecule structure yields a more acceptable value of 137 \AA^2 for the B_P value as well as a positive B_W value, but with an increase in the R value by a factor of two. Here the actual number of water molecules that should be added to the protein appears to be somewhere between 0 and 198. Lattman (1989), as well as other investigators (Otting and Wüthrich, 1989) using NMR experi-

ments, also concluded that at least 4 but no more than 10 water molecules, 6 by NMR, are bound to BPTI. The obvious problem of the number of water molecules, 4–10, bound to BPTI and the exact location of the other 2–6 surface waters does not allow us to further test their methodology using the SAXS profiles of BPTI. However, the results so far indicate that the method of Lattman is still useful for determining whether or not an energy-minimized, three-dimensional structure can be tested by generation of theoretical SAXS profile results and comparing these with experimental data to determine if structures contain bound waters, externally or internally, when they exist in solution.

3.2. Bovine Casein Submicelle structure

We now turn our attention to the comparison of theoretical SAXS profiles of the energy-minimized casein submicelle structures generated by using the Lattman methodology, with published experimental SAXS profiles. Also, we will attempt to ascertain the need for and possible location of bound waters within these structures.

Initially, an extensive search of the literature was employed to acquire the most precise and appropriate SAXS or, small-angle neutron scattering (SANS) profiles for casein under submicellar conditions. Only two papers (Kumosinski *et al.*, 1983; Stothart, 1989) were found in which the data were in graphical form for computer-aided digitization, and the appropriate conditions were used to ensure submicellar (maximally aggregated) casein solution structure. However, one of the papers (Stothart, 1989) contained only SANS data in D₂O and, even though contrast variation experiments were performed at several H₂O/D₂O mixtures, no molecular weights were calculated. Only one report exists whereby precise SAXS experimental profiles were obtained in H₂O and not in D₂O where hydrophobic protein self-association would be increased (Kumosinski *et al.*, 1983) and protein concentrations were such that submicellar casein was present. However, native casein, which contains 10% of α_{s2} -casein (Davies and Law, 1980), was used for this SAXS study and no α_{s2} -casein structure yet exists to be used within the predicted three-dimensional submicellar model of Kumosinski *et al.* (1994). It is presumed a dimer or monomer of α_{s2} -casein could substitute for one

dimer of α_{s1} -casein, or for one monomer of κ -casein respectively.

At this time it would be appropriate to review the details of the previously reported energy-minimized, three-dimensional models of the casein submicelle. In this previous study, two possible submicelle structures consisting of one κ -casein, four α_{s1} -casein and four β -casein monomers were reported; these were called the asymmetric and symmetric structures (Fig. 2). The asymmetric model is shown in Fig. 2A as a ribboned backbone structure with the κ -casein colored cyan while the α_{s1} -casein monomers are red and the four β -casein structures are magenta. Figure 2B shows the backbone model of the symmetric submicelle structure with the same color code as in Fig. 2A. Hence, the only difference between the asymmetric and symmetric submicelle structure depends on the orientation of β -casein dimers within the fundamental synthetic submicelle framework of four α_{s1} - and one κ -casein monomers. Kumosinski *et al.* (1994) have shown via energy minimization using a Kollman force field that such a synthetic submicelle structure is energetically favorable and stabilized mostly by electrostatic and van der Waals interactions with added hydrophobic (side-chain) interactions. This resulting structure contains two cavities consisting of a large number of hydrophobic side chains where β -casein dimers can bind to native casein micelles via hydrophobic interactions (Downey and Murphy, 1970; Rose, 1968). Whether those β -casein dimers are asymmetric with a large dipole moment or symmetric with a lower dipole moment could not be determined. It is the large potential hydrophobic interaction that is responsible for the docking of either β -casein dimer into the framework structure.

Both of the above structures were subjected to the Lattman procedure for comparison with experimental SAXS profiles. The comparisons with experimental results are presented in Fig. 3A. Here, the experimental data are shown as filled triangles with connecting lines while the theoretical curves, using the asymmetric and symmetric structures, are represented by circles and squares, respectively. As can be seen, both structures yield very unfavorable SAX profiles. While agreement with experimental data is moderate at large Q values, unacceptable disagreement occurs at low values; this would yield erroneous molecular weights and radii of gyration. The Lattman

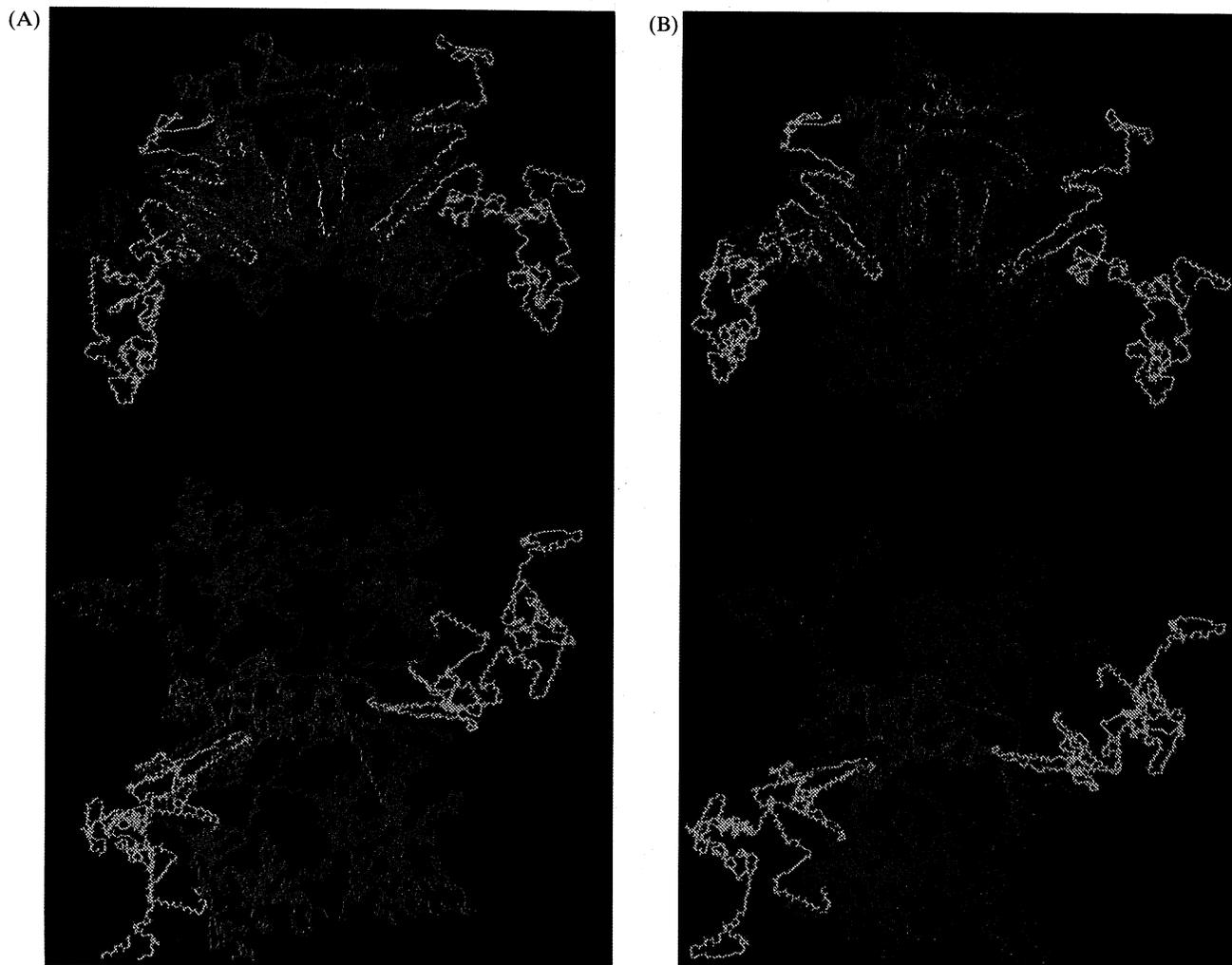


Fig. 2. Refined via energy minimization of submicelle structures, i.e., one κ -casein, four β -casein, and four α_{s1} -casein monomers. (A) ribboned backbone without side chains of asymmetric structure; α_{s1} -casein in red and cyan, κ -casein in blue, and β -casein backbone in magenta (orthogonal view). (B) Backbone structure of symmetric structure; κ -casein in blue, α_{s1} -casein in red and cyan, β -casein in magenta (orthogonal view).

temperature parameters (B values) from these calculations are given in Table III. The B values for the protein are positive with close values of 33 and 35 \AA^2 , respectively, but the large R values of 16.9 and 16.4 for both the asymmetric and symmetric models reflect the poor agreement between theoretical and experimental SAXS profiles. The poor agreement with actual data may reflect the absence of water molecules within these submicellar models, just as four internally bound water molecules were necessary to obtain the lowest R value for the BPTI structure.

With this in mind, 120 water molecules were energy-minimized and docked within the κ -casein

cavity for each of the asymmetric and symmetric submicellar structures (Figs. 2A, B, respectively). This cavity within the κ -casein molecule could in reality contain either bound or free water. The choice of bound water is a reasonable assumption since 120 bound water molecules would mimic the amount of bound water determined via DNMR relaxation results for submicellar casein (Farrell *et al.*, 1989, 1990), i.e., 0.007 g water/g protein, assuming all bound water fell within this cavity.

Subjecting these low-hydrated complexes (120 water molecules to one submicelle molecule) to the Lattman method yielded better results with acceptable R values of 0.682 and 0.611 for the

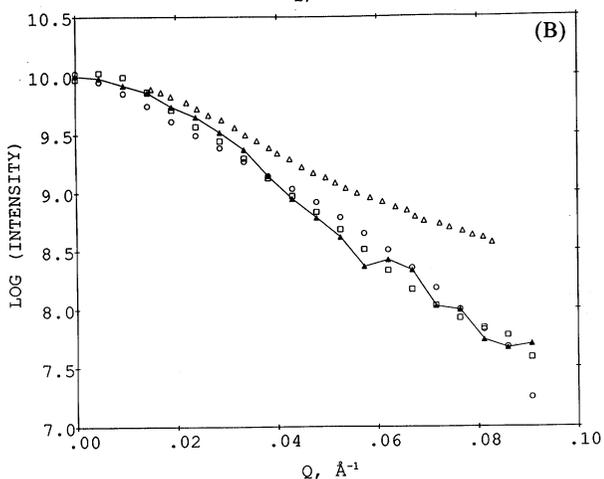
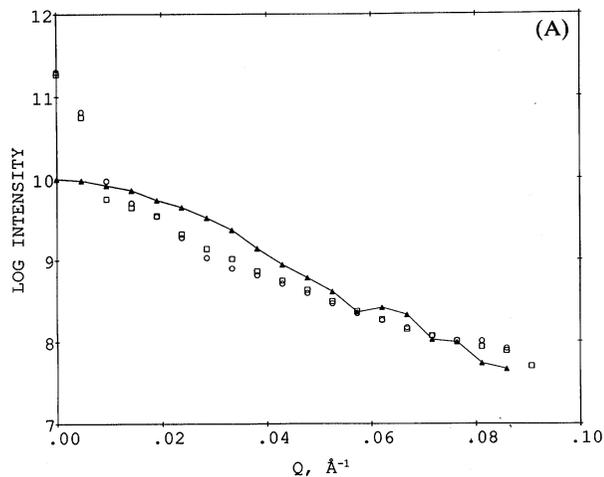


Fig. 3. SAXS profiles of submicellar casein; solid triangles (\blacktriangle) with connected lines represent experimental data (Kumosinski *et al.*, 1983). (A) Theoretical optimized curves from unhydrated asymmetric (circles) and symmetric (squares) energy-minimized structures. (B) Theoretical optimized curves of energy-minimized low-hydration (120 waters) asymmetric (circles) and symmetric (squares) structures; open triangles (\triangle) are experimental small-angle neutron scattering in D_2O (Stothart, 1989). (C) Theoretical optimized curves for energy-minimized high-hydration (2823 waters) asymmetric (circles) and symmetric (squares) structures.

asymmetric and symmetric models, respectively (Table III). Also, the corresponding B_P values are both positive, i.e., 43 \AA^2 , whereas the B_W are both negative, -540 \AA^2 . The negative B_W values could be an artifact of the approximation by Lattman for positioning of the free water molecules at the same position as the protein and bound water atoms.

To follow the same logic as with the BPTI study, we added, using the droplet algorithm, a larger amount of water, i.e., 2073 water molecules to each of the submicelle structures already containing

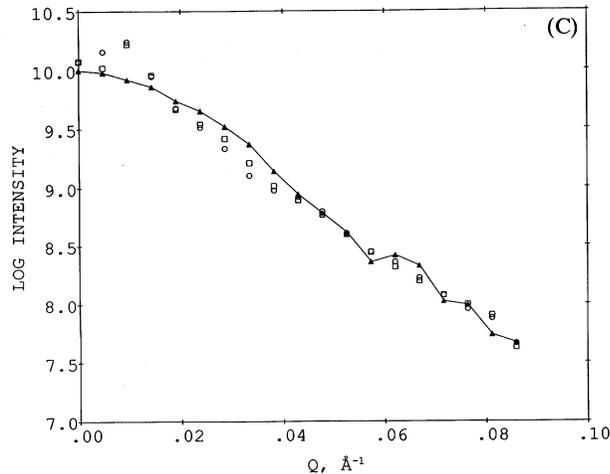


Fig. 3. Continued.

120 water molecules. This amount of water, 2823 mol water/mol protein, would represent a hydration value of 0.244 g water/g protein (acceptable for globular protein). The normal hydrodynamic hydration value for casein submicelles is, however, on the order of 2–3 g water/g protein (up to 40,000 water molecules/submicelle (Farrell *et al.*, 1989; Kumosinski *et al.*, 1983); the gravimetric hydration of isoelectric casein is about 0.7 g water/g protein (8500 water molecules/submicelle) (Farrell *et al.*, 1989, 1990). However, due to the large number of water molecules involved to achieve these values as well as the length of the calculation, it would seem prudent to first attempt to solve the 2823-water-molecule structure to ascertain if any improvement is observed in the R values. Using the Lattman procedure on the two

Table III. Temperature Factors for Submicelle Structures from SAXS^a

Structure	Water	B_P (\AA^2)	B_W (\AA^2)	B_B (\AA)	R
Asymmetric	None	33	-181	—	16.9
Asymmetric	Low	43	-540	413	0.682
Asymmetric	High	36	53	98	1.17
Symmetric	None	35	25	—	16.4
Symmetric	Low	43	-540	503	0.611
Symmetric	High	36	57	130	1.26
Asymmetric ^b	Low	46	-550	436	0.684
Asymmetric ^b	High	30	85	156	1.81
Symmetric ^b	Low	45	-550	500	0.612
Symmetric ^b	High	31	27	100	1.29

^a Parameters defined in Table I.

^b Energy-minimized submicelle-water complex.

droplet structures (i.e., asymmetric and symmetric + 283 water molecules), which we shall define as high-hydration structures, yielded R values that were twice as large as those obtained when using the low-hydration (120 water molecules) structures (see Table III). All hydrated structures had R values 16–20 times lower than the unhydrated structures. B_P values for the asymmetric and symmetric high-hydration structures were 36 \AA^2 and in close agreement with the value of 43 \AA^2 obtained using the low-hydration structure. However, as in the BPTI study, the B_W values were now positive and the B_B values were more realistic, i.e., on the order of 100 \AA^2 . Hence, it would appear that the true hydration value for the submicellar casein structure lies somewhere between the low value of 120 and the higher values of 2823 molecules of water per submicelle. The exact amount and location of these bound waters cannot be determined at this time. Further studies may, in time, resolve this problem.

We next energy-minimized all four hydrated structures, i.e., the low- and high-hydrated asymmetric and symmetric models, to follow the BPTI study. The energetic results of such calculations are given in Tables IV and V for the asymmetric and symmetric structures, respectively. As can readily be observed, the lowest total energy occurs with the high-hydration structures for both models (Tables IV and V). In fact, the high-hydrated asymmetric structure yields a much lower energy value of $-117,949.8 \text{ kcal/mol}$ (Table IV) than the corresponding symmetric structure ($-50,405.6 \text{ kcal/mol}$, Table V). Determination of the water-protein

Table IV. Energy of Hydrated Asymmetric Submicelle Model

	Energy at given hydration level (kcal/mol)		
	None	Low	High
	Bond stretching	208.7	6691.2
Angle bending	4909.0	3884.9	3846.6
Torsional	2868.4	3160.4	3381.8
Out-of-plane bending	200.5	440.3	510.3
1-4 Van der Waals	2697.3	2896.6	2895.5
Van der Waals	-8379.8	-5257.4	33489.6
1-4 Electrostatic	19000.7	19262.9	19326.6
Electrostatic	-43813.0	-54278.8	-236651.5
H bond	-513.9	-508.8	-585.6
Total energy	-22802.1	-23708.6	-117949.8

Table V. Energy of Hydrated Symmetric Submicelle Model

	Energy at given hydration level (kcal/mol)		
	None	Low	High
	Bond stretching	363.2	4606.7
Angle bending	3430.1	3580.9	3581.0
Torsional	3272.6	3478.6	3478.7
Out-of-plane bending	496.1	469.5	469.5
1-4 Van der Waals	2939.1	2965.4	2965.4
Van der Waals	-8200.5	-6737.5	-7700.7
1-4 Electrostatic	19292.8	19326.3	19326.3
Electrostatic	-45120.4	-58493.7	-76836.3
H bond	-519.0	-551.9	-617.7
Total energy	-24047.0	-31355.6	-50405.6

interaction energy for all structures, calculated as the difference between the energy of the hydrated complex and the sum of the energies of the uncomplexed protein and the energy of the water molecules (Table VI), also shows in contrast a lower energy for the symmetric high-hydration model than the corresponding asymmetric model. However, the lower negative values for the high-hydration models over the low-hydrated structures occurs primarily due to the large number of added water molecules. Dividing all interaction energies by the number of water molecules added

Table VI. Hydration Interaction Energy (in kcal/mol) for Refined Casein Submicelle Structures^a

	Asymmetric		Symmetric	
	Low hydration	High hydration	Low hydration	High hydration
	Bond stretching	3933.3	3798.3	19.4
Angle bending	-1024.1	-1062.4	150.8	150.9
Torsional	272.0	493.4	206.0	206.1
Out-of-plane bending	239.8	309.8	-26.3	-26.6
1-4 Van der Waals	1162.6	57.1	-695.8	-1776.8
Van der Waals	262.2	325.9	69.5	33.5
1-4 Electrostatic	-1348.7	-7088.7	-42.5	-2231.9
Electrostatic	5.1	-71.7	-32.9	-98.7
H bond				
Interaction Energy/water molecule added	3701.7	-3040.1	-325.5	-3697.9
	30.9	-0.9	-2.7	-1.3

^a Obtained by subtracting the calculated energy due to the water molecules plus the energy of the nonhydrated submicelles from the energy of the hydrated complexes.

per submicelle, which normalizes the water protein interaction energy, yields a different result (last row of Table VI). In Table VI, it can be easily seen that the low-hydrated symmetric structure yields the lowest comparable energy, i.e., -2.7 kcal/mol per water molecule added to the submicelle. Therefore, from a comparative energetic standpoint, it appears that the low-hydrated symmetric submicelle structure would be stabilized in solution to a higher degree than the asymmetric model, which displays a positive value (Table VI). However, experimental evidence for this hypothesis must still be obtained.

With this in mind, we have subjected all four hydrated refined via energy minimization structures to the Lattman procedure, and the theoretical and experimental profiles for the low- and high-hydrated structures are shown in Figs 3B, C, respectively. In both figures the experimental data are given as filled triangles with connecting lines, while the SAXS profiles that form the asymmetric and symmetric models are circles and squares, respectively. Figure 3C shows clearly that an unacceptable fit to the experimental data is obtained with both high-hydrated structures, especially at low Q values. The zero Q value from which the molecular weight is calculated would yield acceptable agreement with the experimental value, but all other Q values deviate significantly in first a positive and then a negative manner to a Q value of 0.05 \AA^{-1} . Such a nonmonotonic curve with a maximum at low Q values is indicative of high virial or ordering effects. Since this ordering is not observed in the experimental data, it appears that the droplet water added to the low-hydrated structures is not tightly bound water, but is most likely free water molecules and easily exchanged with the bulk solvent molecules.

Figure 3B shows good agreement with theoretical and experimental SAXS profiles. Here, it can readily be seen that the circles are closer to the filled triangles than the squares, which represent the SAXS profile calculated from the low-hydrated asymmetric structure at almost all Q values. This fit is further reflected in Table III by a lower R value of 0.612 for the low-hydrated symmetric structure than for the corresponding asymmetric model, 0.684. However, these results should indicate only that the low-hydrated symmetric structure has only a slightly higher probability of existence in solution. These calculations and experiments do not justify the elimination of a symmetric form. What also can be concluded from

this study is that the casein submicelle structure may be somewhat more rigid than globular proteins. This is seen by the fact that the B_P values for all submicellar casein models yield values on the order of $35\text{--}40 \text{ \AA}^2$ (Table III), while the BPTI values were much higher, i.e., $135\text{--}200 \text{ \AA}^2$ (see Table I). These consistently lower B_P values for casein may be ascribed to the existence of proline residues throughout the polypeptide chain which would yield an open but more rigid backbone structure. Clearly the caseins are not random coils nor are they globular.

In addition, the reported neutron scattering data of casein in D_2O (Stohtart, 1989) is shown in Fig. 3B as open triangles. Here, the neutron data were normalized at zero Q value to be compared with the SAXS profiles. As can easily be observed, large differences exist between experimental SAXS profiles and neutron scattering data in D_2O . Whether this difference is a direct result of structural changes induced by the addition of D_2O is not clear at this time. However, these results do suggest that care must be taken in the use of D_2O in casein solutions.

Finally, because of the apparent rigid nature of all submicellar structures, it would be prudent to ascertain if the protein structure had an influence on the structure of the water molecules within the various hydrated forms. Figure 4A shows these 120 added water molecules within the κ -casein cavity for the energy-minimized low-hydrated submicelle structure. Only the water, shown as "wire-frame" structure, and the ribboned backbone of κ -casein are displayed. The dashed lines indicate the presence of all hydrogen bonds. It can be easily seen that a matrixlike structure of the 120 water molecules is present within the energy-minimized κ -casein cavity. This superstructure of water molecules is obviously due to the influence of protein electrostatic interactions and resembles a solid distorted cylinder as seen by the space-filling model of these waters shown in Fig. 4B. The fact that the κ -casein molecule exhibits a dipolar character is easily observed in Fig. 4C, where an isopotential electrostatic map for the κ -casein is shown. That this protein dipolar character is imposed on the internal water molecules is also evident from the same figure, where the water molecules and their isopotential map are shown above the κ -casein representation. Here, then, is a clear representation of the influence of protein structure and energetics on internal bound water

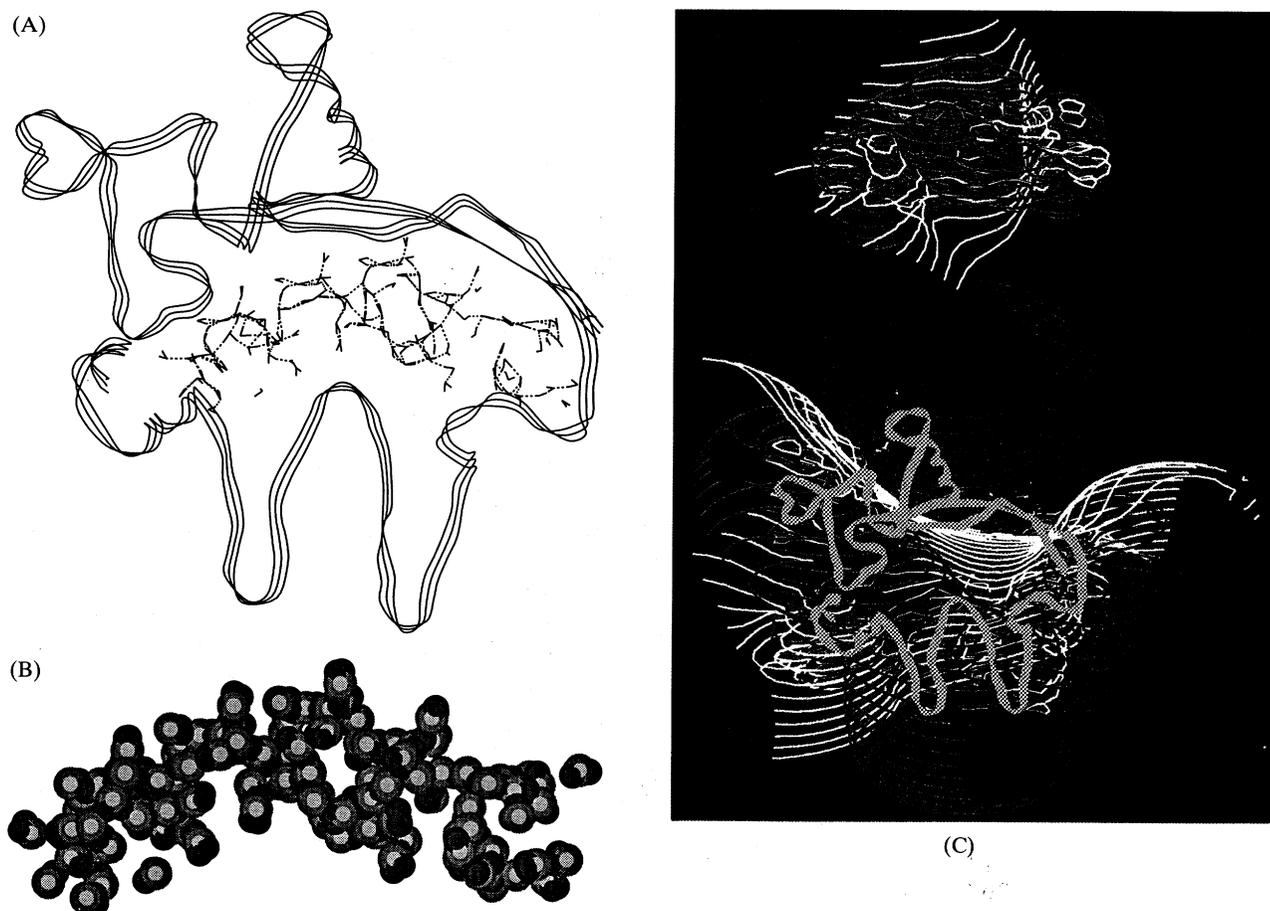


Fig. 4. (A) Energy-minimized docked low-hydration waters (120 waters) displayed with ribboned κ -casein backbone structure; water in black v-shape, with clashed lines representing hydrogen bonding. (B) Space-filled energy-minimized model of low-hydration waters colored by atom types; oxygen in light to medium shading and hydrogen in dark to black shading. (C) Electrostatic isopotential map of (A) with docked waters above κ -casein white-ribboned backbone structure; contoured at 6 Å from the surface structure; +1 kcal in red, -1 kcal in blue, and 0 kcal in yellow.

structures. That the protein structure did not change when the water was added and the complex was energy-minimized further validates the apparent rigidity of the casein backbone structure.

As an afterthought, the resulting energy-minimized high-hydration asymmetric structure is presented in Fig. 5. Here, 2703 water molecules were added to the low-hydrated form using the Sybyl droplet algorithm. As can be observed, these added waters, given in cyan, lie within small channels and fissures caused by the docking of the α_{s1} -casein monomers and β -casein monomers to the κ -casein. These solvent exchangeable waters may still be present when casein is dried by lyophilization or some other processing method. Hence, care during a drying process would be

prudent, since loss of this water may disrupt the integrity of the submicellar structure, which, in turn, ultimately determines the functional properties of the caseinate.

As emphasized in previous papers on the monomeric caseins and in Kumosinski (1994) on construction of the submicelle, it must be kept in mind that these structures represent working models. They are not the final native structures, but are presented to stimulate discussion and to be modified as future research unravels the nature of these noncrystallizable proteins. Inspection of a recent drawing of the casein micelle by Holt (1992) demonstrates how structures such as those presented here could be further aggregated into the casein micelle. Continued dialogue and research in

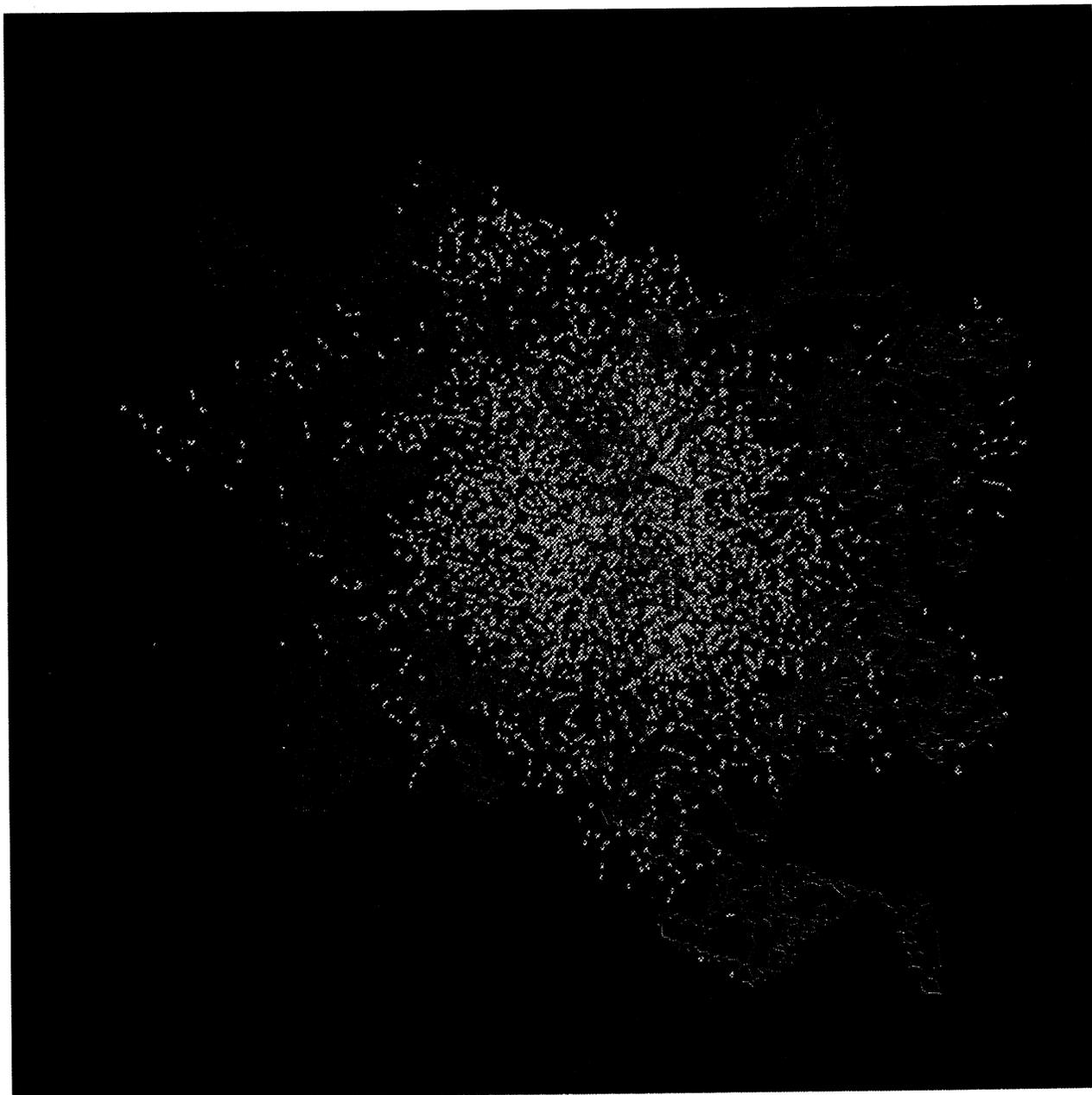


Fig. 5. Backbone asymmetric structure of casein submicelle with waters from droplet algorithm i.e., 2823 waters; κ -casein in blue, α_{s1} -casein in red, β -casein in magenta, oxygen from droplet waters in cyan.

this area may bring together the new concepts necessary to finally produce an accurate model. It is hoped that this work is a start in that direction.

REFERENCES

- Clore, G. M., Bax, A., Wingfield, P. T., and Gronenborn, A. M. (1990). *Biochemistry* **29**, 5671–5676.
- Davies, D. T., and Law, A. J. R. (1980). *J. Dairy Res.* **47**, 83–90.
- Downey, W. K., and Murphy, R. F. (1970). *J. Dairy Res.* **37**, 361–372.
- Edsall, J. T., and McKenzie, H. A. (1983). *Adv. Biophys.* **16**, 53–183.
- Farrell, H. M., Jr. (1988). Physical equilibria: Proteins, in *Fundamentals of Dairy Chemistry*, 3rd ed. (Wong, N., ed.), Van Nostrand Reinhold, New York, pp. 461–510.
- Farrell, H. M. Jr., Pessen, H., and Kumosinski, T. F. (1989). *J. Dairy Sci.* **72**, 562–574.

- Farrell, H. M., Jr., Pessen, H., Brown, E. M., and Kumosinski, T. F. (1990). *J. Dairy Sci.* **73**, 3592–3601.
- Gelin, B. R., and Karplus, M. (1979). *Biochemistry* **18**, 1256–1259.
- Holt, C. (1992). *Adv. Protein Chem.* **43**, 63–113.
- Kumosinski, T. F., and Farrell, H. M., Jr. (1991). *J. Protein Chem.* **10**, 3–11.
- Kumosinski, T. F., Pessen, H., Farrell, H. M., Jr., and Brumberger, H. (1983). *Arch. Biochem. Biophys.* **266**, 548–561.
- Kumosinski, T. F., King, G., and Farrell, H. M., Jr. (1994). *J. Protein Chem.*, this issue.
- Lattman, E. E. (1989). *Proteins: Struct. Funct. Genet.* **5**, 149–155.
- Otting, G., and Wüthrich, K. (1989). *J. Am. Chem. Soc.* **111**, 1871–1875.
- Pessen, H., Kumosinski, T. F., and Farrell, H. J., Jr. (1988). *J. Ind. Microbiol.*, **3**, 89–104.
- Pickover, C. A., and Engleman, D. M. (1982). *Biopolymers* **21**, 817–832.
- Rose, D. (1968). *J. Dairy Sci.* **51**, 1897–1902.
- Schmidt, D. G. (1982). In *Developments in Dairy Chemistry I* (Fox, P. F., ed.), Applied Science, Essex, England, pp. 61–93.
- Stothart, P. H. (1989). *J. Mol. Biol.* **208**, 635–639.