PREDICTION OF THE CONFORMATION OF THE COW AND SHEEP κ-CASEINS

MARIE-HENRIETTE LOUCHEUX-LEFEBVRE, JEAN-PIERRE AUBERT, AND PIERRE JOLLÈS, *Institut de Recherches sur le Cancer de Lille, Lille, and Laboratoire des Protéines, Université de Paris V, Paris, France.*

ABSTRACT The secondary structures of cow and sheep κ-caseins were established according to the predictive rules of Chou and Fasman. The diagrams derived from this treatment allowed us to study the chymosin-sensitive bond (milk-clotting process), as well as the glycosylation and phosphorylation sites, found to be situated in β-turns. Despite a high variability between the primary structures of the COOH-terminal part (caseinoglycopeptide) of cow, sheep, and also other caseins, the secondary structures of the biologically important sites were found to be conserved.

INTRODUCTION

To establish whether the knowledge of the secondary structure might be a useful complement to primary sequence data in the study of biologically important sites, we attempted to predict the secondary structures of κ-caseins of different species. During recent years, a number of various predictive algorithms have been published that predicted with varying degrees of success the secondary structure of a protein from its sequence data. Such algorithms are based on the known amino acid sequences and the secondary structures delineated by the crystallographic studies. Argos et al. (1976) combined five methods to a "joint prediction histogram" and stated that the predictions based on this histogram should be nearly equivalent to any individual prediction. In this work, we used the method of Chou and Fasman (1974), the predictive method currently used. The secondary structures of cow and sheep κ-caseins were predicted, as the corresponding primary structures were already known.

MATERIALS AND METHODS

κ_A-Caseins

Cow κ_A-casein was prepared according to McKenzie and Wake (1961) and sheep κ_A-casein according to Alais and Jollès (1967).

Prediction of Secondary Structure

α HELIX AND β-SHEET REGIONS For the conformational assignments of the amino acids we used the values indicated by Fasman et al. (1976). These parameters were slightly different from those proposed by Chou and Fasman in their first paper (1974) concerning the pre-

Inquiries should be addressed to P. Jollès, Université de Paris V.

dictive determination of secondary structures of proteins, and were re-evaluated from the X-ray data of 29 proteins. In Fig. 1 the assignments of helical (α) and β-sheet (β) potentials are given under each residue: *H*, strong helix of β-sheet-former, *h*, helix of β-sheet-former, *I*, weak helix or β-sheet-former, *i*, helix or β-sheet indifferent, *b*, helix or β-sheet-breaker, and *B*, strong helix or β-sheet-breaker. A conformational assignment was given to each residue and

Sheep	1	Q	Q	Q	N	Q	E	E	R	I	C	C	E	K	D	E	R	F	F	D	D
Cow	1		E					Q	P		R									S	
	α	h	h/H	h	b	h	H	H/h	i/B	h	b/i	b	H	h	I	H	i	h	h	I/i	I
	β	h	h/B	h	i	h	B	B/h	i/b	H	h/i	h	B	b	b	B	i	h	h	b/b	b
Sheep	21	K	I	A	K	Y	I	P	I	Q	Y	V	L	S	R	Y	P	S	Y	G	L
Cow	21																				
	α	h	h	H	h	b	h	B	h	h	b	h	H	i	i	b	B	i	b	B	H
	β	b	H	i	b	H	H	b	H	h	H	H	h	b	i	H	b	b	H	b	h
Sheep	41	N	Y	Y	Q	Q	R	P	V	A	L	I	N	N	Q	F	L	P	Y	P	Y
Cow	41						K														
	α	b	b	b	h	h	i/h	B	h	H	H	h	b	b	h	h	H	B	b	B	b
	β	i	H	H	h	h	i/b	b	H	i	h	H	i	i	h	h	h	b	H	b	H
Sheep	61	Y	A	K	P	V	A	V	R	S	P	A	Q	T	L	Q	W	Q	V	L	P
Cow	61					A								I							S
	α	b	H	h	B	h/H	H	h	i	i	B	H	h	i/h	H	h	h	h	h	H	B/i
	β	H	i	b	b	H/i	i	H	i	b	b	i	h	h/H	H	h	h	h	H	h	b/b
Sheep	81	N	A	V	P	A	K	S	C	Q	D	Q	P	T	A	M	A	R	H	P	H
Cow	81		T								A				T						
	α	b	H/i	h	B	H	h	i	b	h	I/H	h	B	i	H/i	H	H	i	I	B	I
	β	i	i/h	H	b	i	b	b	h	h	b/i	h	b	h	i/h	h	i	i	i	b	i
Sheep	101	P	H	L	S	F ↓	M	A	I	P	P	K	K	D	Q	D	K	T	E	I	P
Cow	101													N							
	α	B	I	H	i	h	H	H	h	B	B	h	h	I/b	h	I	h	i	H	h	B
	β	b	i	h	b	h	h	i	H	b	b	b	b	b/i	h	b	b	h	B	H	b
Sheep	121	A	I	N	T	I	A	S	A	E	P	T	V	H	S	T	P	T	T	E	A
Cow	121	T							G				∘	∘							
	α	H/i	h	b	i	h	H	i	H/B	H	B	i	h	I	i	i	B	i	i	H	H
	β	i/h	H	i	h	H	i	b	h/b	B	b	h	H	i	b	h	b	h	h	B	i
Sheep	141	V	V	N	A	V	D	N	P	E	A	S	S	E	S	I	A	S	A	P	E
Cow	139		E	S	T		A	T	L		D		P		V		E		P		
	α	h	h/H	b/i	H/i	h	I/H	b/i	B/H	H	H/I	i	i/B	H	i/H	h	H/H	i	H/B	B	H
	β	H	H/B	i/b	i/h	H	b/i	i/h	b/h	B	h/b	b	b/b	B	b/H	H	i/B	b	i/b	b	B
Sheep	161	T	N	T	A	Q	V	T	S	T	E	V									
Cow	159	I			V						A										
	α	i/h	b	i	H/h	h	h	i	i	i	H/H	h									
	β	h/H	i	h	i/H	h	H	h	b	h	B/i	H									

FIGURE 1 Assignments of helical and β-sheet potentials for each residue of sheep and cow κ_A-caseins. Assignments in the first row under each residue refer to helical potential and in the second row to β-sheet potential as defined by Fasman et al. (1976). When a replacement between the two κ_A-caseins occurs, the corresponding two assignments are indicated. The amino acid sequence data were established by Jollès et al. (1974). The one-letter amino acid code was used according to Dayhoff (1976). ∘, deletion; ↓ chymosin-sensitive bond.

the nucleation centers were determined by the predictive rules of Chou and Fasman (1974): "When four helix-formers out of six residues or three β-formers out of five residues are found clustered together in any native protein segment, the nucleation of these secondary structures begins and propagates in both directions until terminated by tetrapeptide breakers with 50% or more indifferent or breaking residues" (Chou et al., 1975).

β-TURNS A β-turn consists of four amino acid residues, in a region where the protein chain folds back on itself by nearly 180°. In a survey of 29 proteins, Chou and Fasman (1977) indicated the frequency of occurrence for a certain residue to adopt the first, second, third, and fourth position of a β-turn, f_i, f_{i+1}, f_{i+2}, and f_{i+3} respectively. The relative probability that a tetrapeptide forms a β-turn is $p_t = f_i \cdot f_{i+1} \cdot f_{i+2} \cdot f_{i+3}$ (Lewis et al., 1973). These authors took $p_t = 1.10^{-4}$ as a cut-off value in predicting the β-turns. If $< P_\alpha >$, $< P_\beta >$, and $< P_T >$ are the structural parameters for α helix, β-sheet, and β-turn respectively, we assert that a tetrapeptide has a high probability of occurrence as a β-turn if the following conditions are met: $< P_\alpha > \ll < P_T >$; $< P_\beta > \ll < P_T >$; $p_t > 1.10^{-4}$.

Circular Dichroism (CD)

The dichroic spectra were recorded on a Jobin-Yvon dichrograph R. J. Mark III, between 180 and 250 nm. A fused quartz cell of 0.1-mm path length was used. The concentrations were determined either by the method of Lowry et al. (1951) or by ultraviolet (UV) spectrometry ($\epsilon_{280} = 1.2$). Generally, the solvent conditions were 0.214 NaF in H_2O. The pH of these unbuffered solutions was 7.3. When necessary, NaF concentration was changed by adding small volumes of a concentrated salt solution. Duplicate runs were carried out. The ellipticities [Θ] were expressed in degree · centimeter squared · decimole^{-1}. The mean residue weight is 111. Considering that the sugar content of κ_A-casein is 5% (Alais and Jollès, 1961), the contribution by the sugars to the dichroic spectrum of the glycoprotein is negligible (Aubert and Loucheux-Lefebvre, 1976). A nonlinear program, nonlinear Gaussian version, was used to resolve the CD curve into Gaussian bands (Thiery, 1969).

RESULTS AND DISCUSSION

The κ-casein molecules consist of two parts separated by the chymosin-sensitive bond (Jollès, 1966; Delfour et al., 1965; see also below): para-κ-casein (NH_2-terminal moiety, residues 1–105) and κ-caseinoglycopeptide (COOH-terminal moiety, residues 106–169 for cow and 106–171 for sheep κ-caseins). The primary structures were previously described by Brignon et al. (1972) and Jollès et al. (1972) for cow para-κ-casein, Mercier et al. (1972) and Jollès et al. (1973) for cow κ-caseinoglycopeptide and by Jollès et al. (1974) for sheep κ-casein (Fig. 1). We wish to discuss the secondary structures of both κ-caseins and of some of their biologically important sites (see below). For a few other κ-caseins, only the structure of their COOH-terminal part was determined; thus we restricted the discussion below concerning the phylogenetic aspects to the caseinoglycopeptides.

Predicted Secondary Structure of Cow κ_A-Casein

The predicted α-helix and β-sheet regions are summarized in Table I. The relative composition of the conformational regions are as follows: 23% α-helix (5 regions in-

TABLE I

CONFORMATIONAL PREDICTIONS CONCERNING COW κ_A-CASEIN CALCULATED FOR α-HELIX, β-SHEET, AND β-TURN REGIONS

	Sequence	$<P_\alpha>$	$<P_\beta>$
α-helix	1–7	1.16	0.86
	12–18	1.17	0.81
	90–97	1.07	0.96
	102–108	1.16	1.11
	137–145	1.16	1.02
β-sheet	22–26	1.08	1.25
	28–32	1.03	1.43
	40–45	0.91	1.22
	48–56	1.07	1.22
	73–80	1.08	1.28
	119–126	0.94	1.18
	159–169	0.97	1.23

β-turn	Tetrapeptide	$<P_\alpha>$	$<P_\beta>$	$<P_T>$	p_t
35–38	Y P S Y	0.68	1.06	1.31	$3.8.10^{-4}$
58–61	Y P Y Y	0.66	1.24	1.23	$3.5.10^{-4}$
69–72	S P A Q	0.97	0.81	1.15	$1.2.10^{-4}$
85–88	A K S C	1.01	0.88	1.07	$1.1.10^{-4}$
98–101	H P H P	0.78	0.71	1.23	$2.7.10^{-4}$
109–112	P P K K	0.86	0.64	1.26	$2.1.10^{-4}$
113–116	D Q D K	1.07	0.73	1.23	$2.4.10^{-4}$
129–132	E P T S	0.92	0.71	1.17	$1.2.10^{-4}$
133–136	T P T T	0.76	1.03	1.10	$1.3.10^{-4}$
149–152	S P E V	0.98	0.84	1.05	$1.5.10^{-4}$

α-helix 23%; β-sheet 31%; β-turns 14%; $<P_\alpha>$, $<P_\beta>$, $<P_T>$ and p_t values were established as indicated in the Methods. For the sequence see Fig. 1.

cluding 39 residues) and 31% β-sheet (7 regions including 52 residues). The computed β-turn profile for cow κ_A-casein is indicated in Fig. 2. There are 21 peaks above the cut-off point $p_t = 1.10^{-4}$, indicating that the corresponding tetrapeptides probably constitute β-turns. However, 11 of the latter could be excluded because they were already predicted to be situated in α-helical or β-sheet regions. 10 β-turns remain possible (Table I), corresponding to 40 residues or 24% of the molecule.

The complete predicted secondary structure of the protein moiety of cow κ_A-casein is shown in a schematic diagram (Fig. 3). The two β-sheet domains, 22–32 and 40–56, seem to occur in an antiparallel β formation, as they are separated by the predicted β-turn 35–38.

For the helical sequences 1–7, 12–18, and 90–97, the following parameters were established: $<P_\alpha> > <P_\beta>$ with $<P_\beta> < 1$. For the sequences 102–108 and 137–145, $<P_\beta> > 1$, with $<P_\alpha> > <P_\beta>$ and therefore these sequences were predicted to be in an α-helical structure. However, for residues 102–108 ($<P_\alpha> = 1.16$ and $<P_\beta> = 1.11$) the α-helical and β-sheet potentials are nearly identical. This part of the molecule and its special behavior will be discussed in detail in the part devoted to the chymosin-sensitive bonds.

Predicted Secondary Structure of Sheep κ_A-Casein: Comparison with Cow κ_A-Casein.

The primary structure of sheep κ_A-casein was established by Jollès et al. (1974). 2 deletions and 29 replacements were noted when the cow and sheep κ_A-casein sequences were compared (Fig. 1); however, the changes are not equally distributed, as discussed below.

The predicted secondary structure for sheep para-κ_A-casein (NH_2-terminal moiety of κ-casein) was the same as that established for cow κ_A-casein; more especially no variation occurred in the four β-turns. 10 out of the 12 replacements noted between both proteins containing 105 residues were conservative. 5 of them (residues 8, 10, 19, 46, and 82) were situated in unordered regions. Residue 8 (Arg) can be included in the α-helical sequence 1–7, as Arg^+ has an I_α assignment near the COOH-terminal end of a helical region. In the 90–97 sequence, the replacements at positions 90 (Ala → Asp) and 94 (Thr → Ala) gave $< P_\alpha >$ = 1.07 instead of $< P_\alpha >$ = 1.09 and therefore this helical region was conserved.

The primary structures of sheep and cow κ_A-caseinoglycopeptide (COOH-terminal moieties of κ-caseins) were significantly different as 17 replacements and 2 insertions were observed for a total of 66 and 64 residues, respectively. The two additional amino

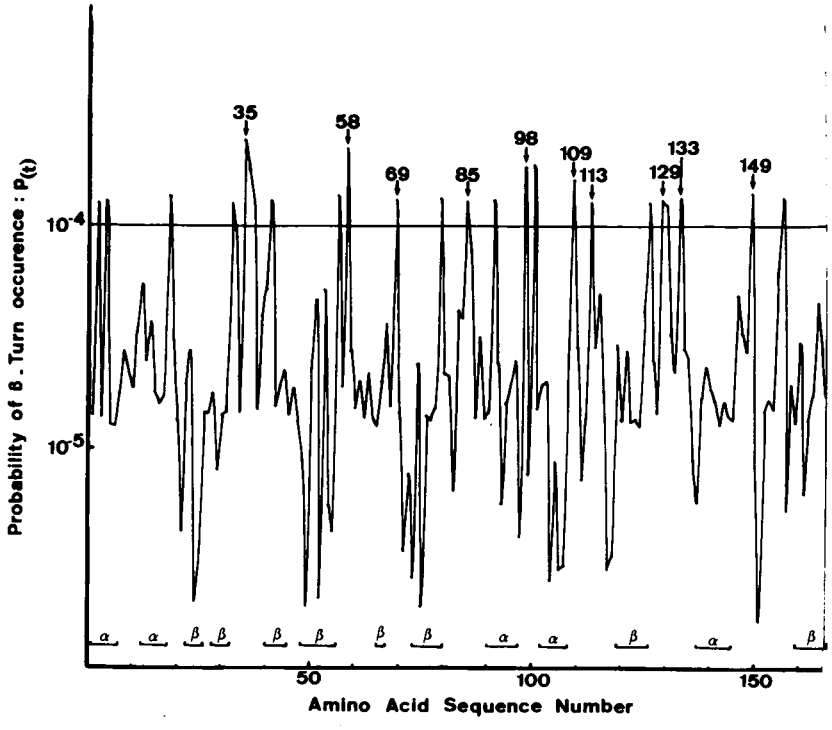

FIGURE 2 Probability of tetrapeptide β-turns in cow κ_A-casein. The horizontal line corresponds to an arbitrary cut-off value of 1.10^{-4}. 11 predicted β-turns were excluded as already belonging to other regions. The α-helical and β-sheet regions determined in Table I are indicated.

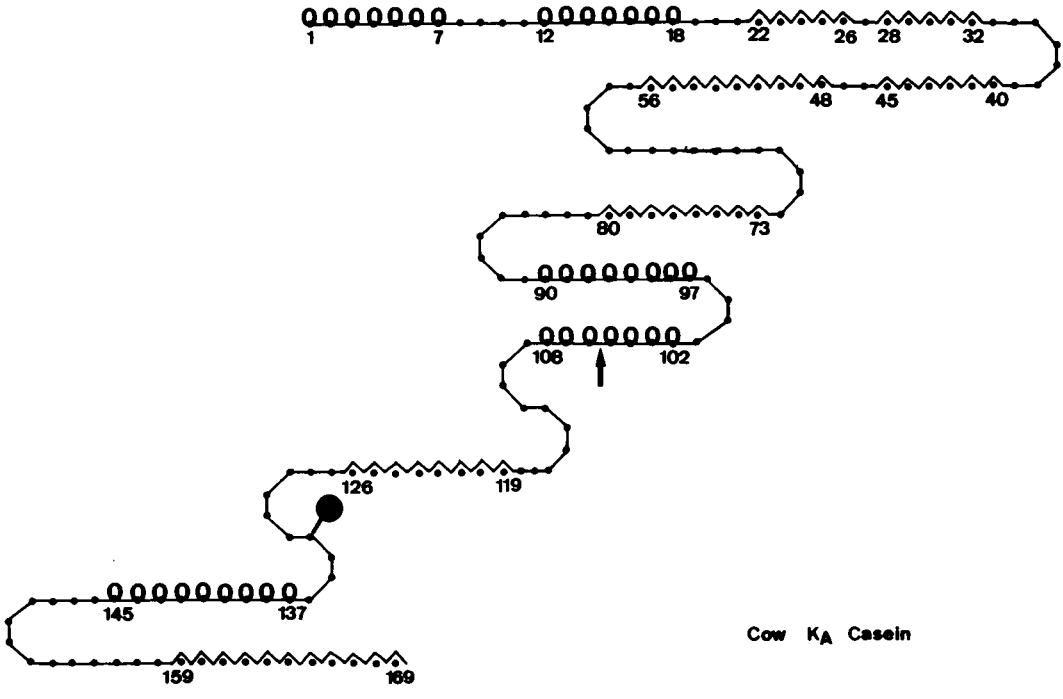

FIGURE 3 Schematic diagram of the predicted secondary structure of cow κ_A-casein. Residues are represented in their respective conformational state: helical (o), β-sheet ($\wedge$), coil (-), β-turn ($\supset$). Thr-133 carries the carbohydrate moiety, the arrow between residues 105 and 106 indicates the chymosin-sensitive bond, and Ser 149 is phosphorylated.

acids occurring in the sheep peptide, Val-His, appeared between residues 131 and 132 of cow κ_A-casein and provoked a structural change in this area. The relative probability of formation of β-turns in sheep κ_A-casein is described in the probability profile shown in Fig. 4, where the helical and β-sheet regions are also mentioned. The corresponding values of $< P_\alpha >$, $< P_\beta >$, $< P_T >$, and p_t are indicated in Table II. The predicted secondary structure of sheep κ_A-casein is drawn in Fig. 5. In its COOH-terminal (caseinoglycopeptide) part was situated one important variation when compared to the corresponding cow sequence. In the region 139–163 the replacements were particularly numerous and a β-turn with a high potentiality was observed in the area 151–154 as $< P_T > = 1.28$ and $p_t = 1.36 \cdot 10^{-4}$.

The location of the sugar part as well as the region containing the chymosin-sensitive bond will be discussed in a separate section of this paper.

CD Studies

The CD spectrum of cow κ_A-casein at 23°C in 0.214 M NaF is shown Fig. 6. It exhibits a negative band centered at about 198 nm. The resolution of this curve into Gaussian bands yields four negative maxima at 222, 214, 207, and 198 nm with $[\theta]$ values of $-4{,}100$, $-2{,}900$, $-4{,}900$ and $-24{,}800$, respectively. The α-helix content was

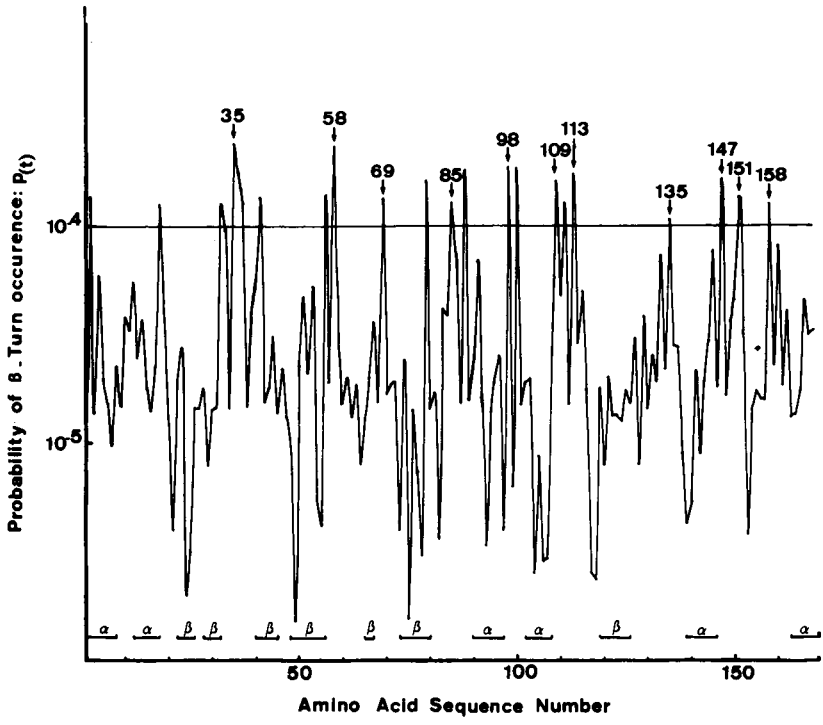

FIGURE 4 Probability of tetrapeptide β-turns in sheep κ_A-casein. The horizontal line corresponds to an arbitrary cut-off value of 1.10^{-4}. Nine predicted β-turns were excluded as already belonging to other regions.

TABLE II

CONFORMATIONAL PREDICTIONS CONCERNING SHEEP κ_A-CASEINOGLYCOPEPTIDE, THE COOH-TERMINAL MOIETY OF SHEEP κ_A-*CASEIN* RELEASED BY CHYMOSIN

		Sequence	$<P_\alpha>$	$<P_\beta>$	
α-helix		102–108	1.16	1.11	
		139–146	1.15	1.07	
β-sheet		119–126	1.02	1.14	
		164–171	1.07	1.10	
β-turn	Tetrapeptide	$<P_\alpha>$	$<P_\beta>$	$<P_T>$	p_t
109–112	P P K K	0.86	0.65	1.26	$2.1.10^{-4}$
113–116	D Q D K	1.07	0.73	1.23	$2.4.10^{-4}$
135–138	T P T T	0.76	1.03	1.13	$1.3.10^{-4}$
147–150	N P E A	1.04	0.66	1.12	$2.2.10^{-4}$
151–154	S S E S	0.95	0.65	1.28	$1.4.10^{-4}$
158–161	A P E T	1.08	0.70	0.97	$1.1.10^{-4}$

$<P_\alpha>$ and $<P_\beta>$ correspond to the 102–108 sequence (see Fig. 1). For details, see the Methods and Fig. 1.

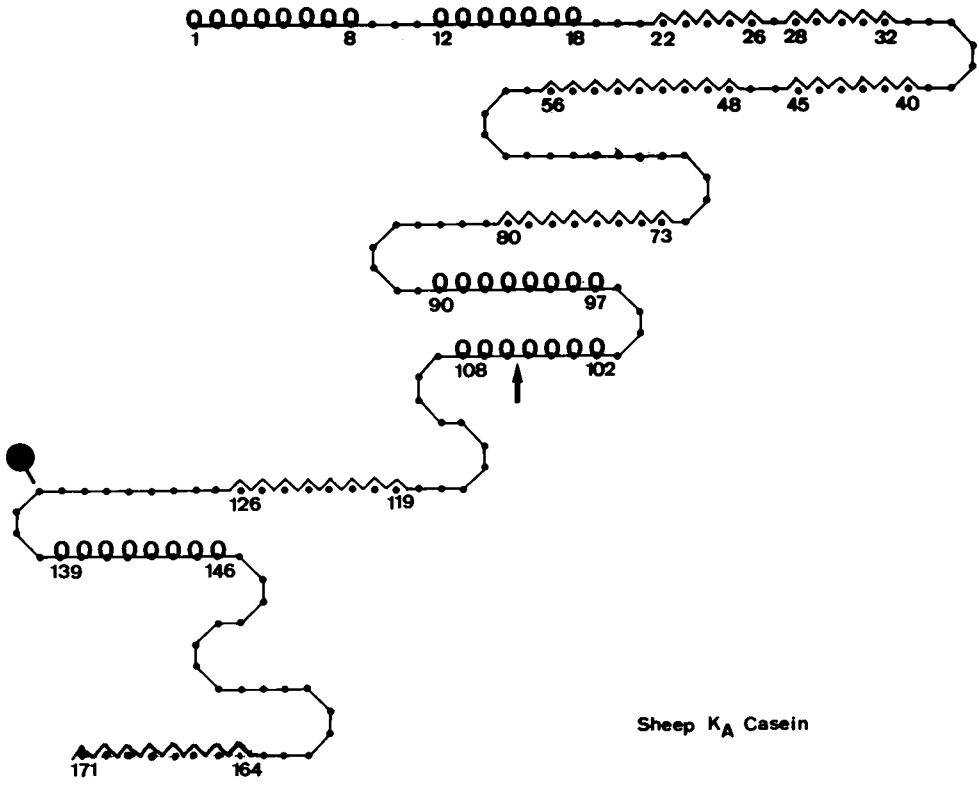

FIGURE 5 Schematic diagram of the predicted secondary structure of sheep κ_A-casein. The symbols have the same meaning as in Fig. 3.

then estimated from the $n \rightarrow \Pi^*$ transition, i.e. the band centered at 222 nm according to Chen et al. (1974), taking the value of $-30{,}000$ for the helix as standard. The amount of β structure was estimated similarly, using the band at 214 nm and taking $-9{,}200$ for the β-standard. By these procedures, the helix and β-sheet contents, 14% and 31% respectively, were thus established.

These results were compared to the predicted conformation. The estimated β-sheet percentage was in excellent agreement with the predicted conformation. A less satisfactory value was found for the percentage of α-helix: the α-helix prediction was higher than that found by circular dichroism. To explain this discrepancy it is interesting to note that the 102–108 helical sequence has about the same $< P_\alpha >$ and $< P_\beta >$ parameters; therefore its structure in solution should depend on different physical parameters, such as the concentration, the temperature, the solvent, etc. Moreover sequence 90–97 presents a weak $< P_\alpha > = 1.07$ value due to the presence of an α-helical breaker ($H_3 hi_3 B$), which decreases the stability of this sequence.

To check the stability of the secondary structure, a study of CD with temperature was carried out; because the hydrophobic forces increase with temperature, the dichroic spectrum of κ_A-casein was recorded at different temperatures, and Fig. 7 shows the variation of the ellipticities at 222, 214 and 198 nm. $[\Theta]_{222}$ and $[\Theta]_{214}$ increase with

temperature and more especially the helicity varies from 14% at 23°C to 20% at 90°C; the latter value agrees rather well with that from the predictive determination.

The quantitative determination of the unordered structure is inaccurate, as the band centered at 198 nm represents the contribution of both unordered and β-turns structures. But the variation of the ellipticity at 198 nm with the temperature shows that

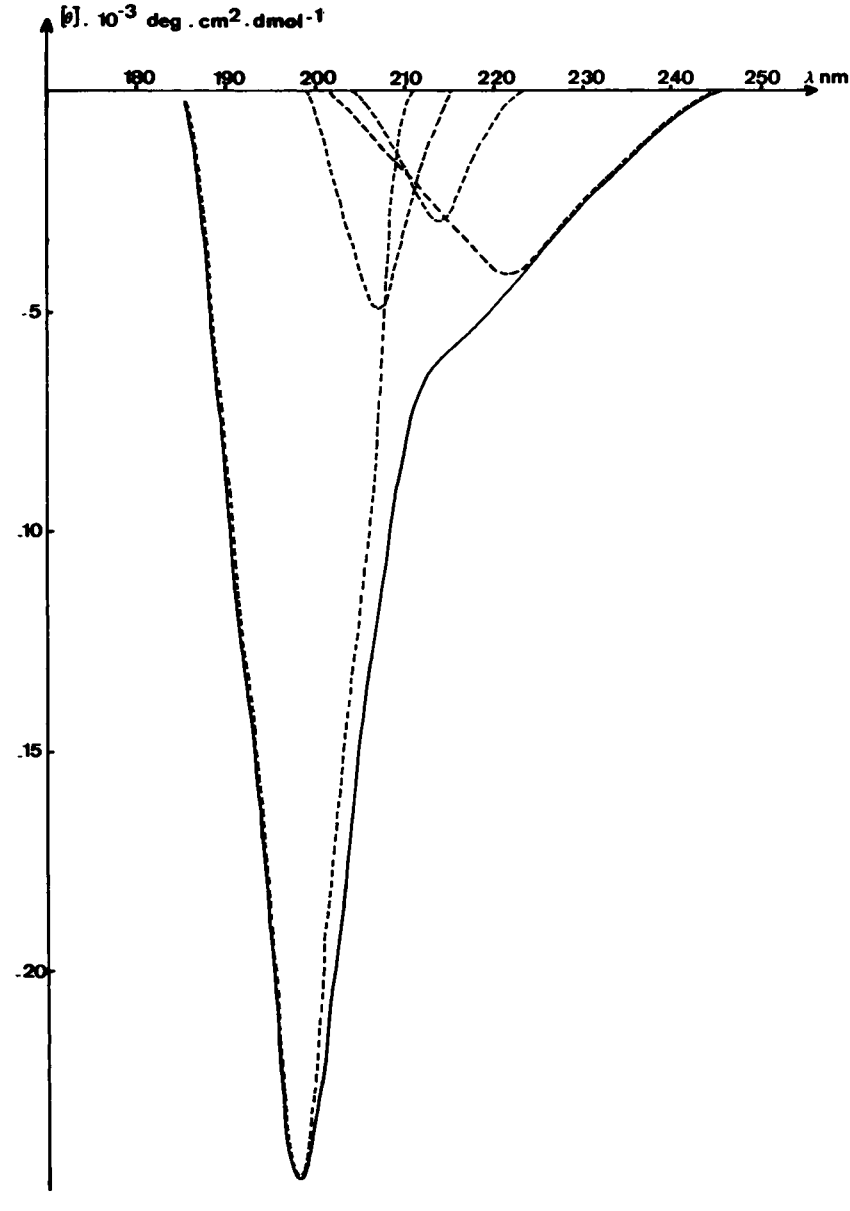

FIGURE 6 Circular dichroic spectrum of cow κ_A-casein at 23°C in solution in 0.214 M NaF. Its resolution into Gaussian curves centered at 222, 214, 207, and 198 nm is shown (- - -).

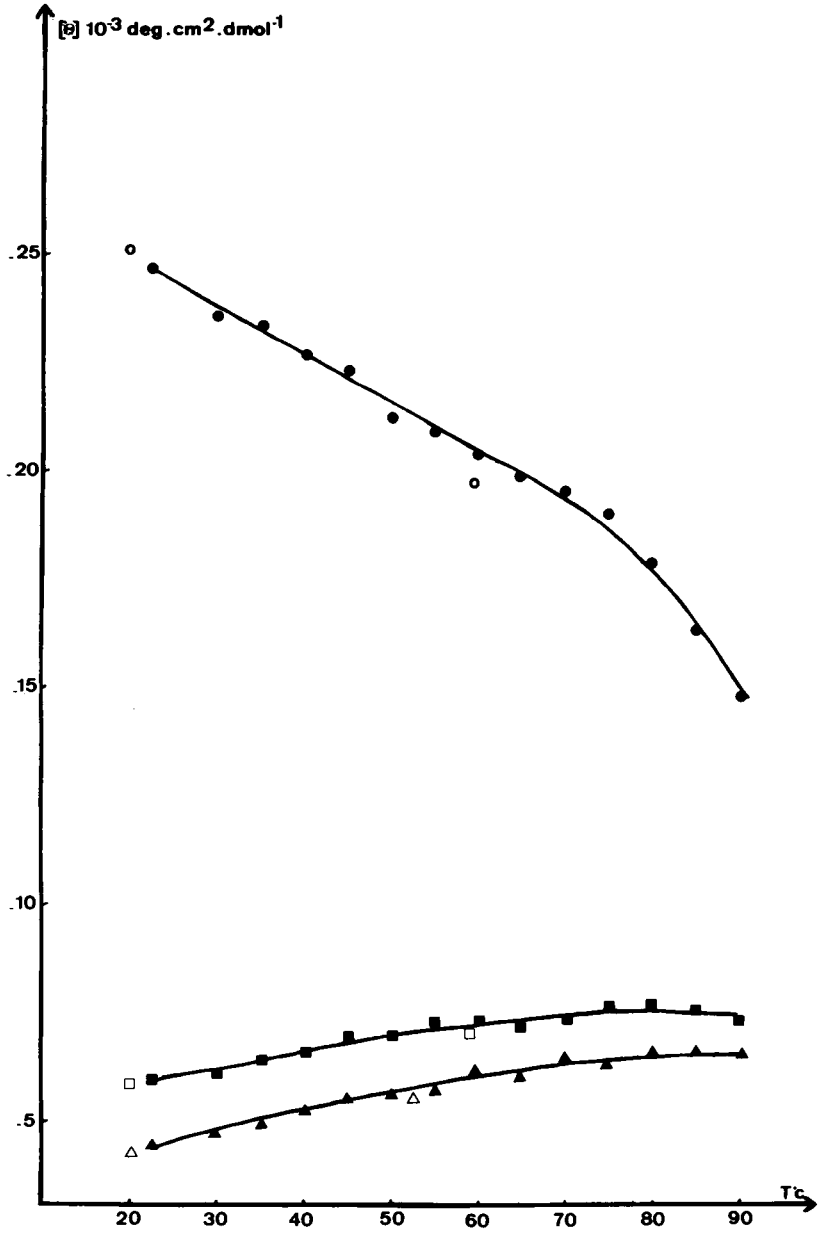

FIGURE 7 Variation of the ellipticities of cow κ_A-casein with temperature: $[\Theta]_{222}$ (▲), $[\Theta]_{214}$ (■), and $[\Theta]_{198}$ (●). △, □, ○ correspond to the ellipticities at 222, 214, and 198 nm, respectively, after denaturation.

$[\theta]_{198}$ decreases when the temperature increases, as expected from the variation of $[\theta]_{222}$.

The Carbohydrate-Peptide Linkage

Previously Aubert et al. (1976) found that 37 carbohydrate-peptide linkages, including *O*-glycosidic (*N*-acetylgalactosamine → Ser/Thr) and *N*-glycosidic (*N*-acetylglucosamine → Asn) bonds, characterized in 19 glycoproteins, were situated predominantly in β-turn regions. These results were corroborated by Beeley (1976, 1977) and by Nagarajan and Rao (1977). The same organization was found in cow and sheep κ_A-caseins, where the unique polysaccharide chain was *O*-glycosidically linked to a threonine residue, 133 and 135, respectively (Jollès et al., 1973). Indeed, Figs. 3 and 5 indicate that for both caseins the corresponding threonine residue is situated in a β-turn and, more especially, is the first residue of the tetrapeptide.

The Site of Phosphorylation

Cow and sheep κ_A-caseins are phosphorylated glycoproteins. The site of phosphorylation of cow κ_A-casein was serine residue 149. The site of phosphorylation of sheep κ_A-casein was not yet determined but by comparison with goat κ_A-casein, it seems that serine residue 151 was implicated in the phosphorylation (Mercier et al., 1976). As observed from Figs. 3 and 5, these residues belong to a β-turn. This result was expected, as Small et al. (1977) demonstrated that phosphorylation in proteins occurs with high probability in a β-turn.

In the case of κ_A-caseins, Mercier et al. (1976) proposed a "code sequence" recognized by a specific phosphoryl kinase, which reacts on a serine residue belonging to a Ser-*X*-Glu sequence, in which *X* is any residue of amino acid. Indeed, residues Ser (149) and Ser (151) of cow and sheep caseins belong to the sequences Ser-Pro-Glu and Ser-Ser-Glu, respectively. Another site of phosphorylation should be expected for the serine residue 127 of sheep and cow κ-caseins belonging to the sequences Ser-Gly-Glu and Ser-Ala-Glu, respectively. However, the latter do not adopt a β-turn structure and in fact they are not phosphorylated. In other words, the code of recognition, Ser-*X*-Glu, might be a necessary but not sufficient condition and the presence of a β-turn seems to be an additional one.

A somewhat opposite situation was encountered in goat caseinoglycopeptide, where Mercier et al. (1976) claimed that residue Ser (168) belonging to the sequence Ser-Thr-Glu was phosphorylated. No β-turn could be predicted here; a possible explanation might be that this residue was situated near the COOH-terminal end of the protein and in a sequence with very similar $< P_\alpha >$ and $< P_\beta >$ parameters closely related to 1.

Phylogenetic Aspects

Only the primary structures of cow and sheep κ-caseins are so far known. However, the amino acid sequences of the COOH-terminal moieties (caseinoglycopeptides) of some other κ-caseins have been established (Mercier et al., 1976; Jollès, 1975). We were thus able to compare the secondary structures of the goat, sheep, cow, zebu, and water buffalo caseinoglycopeptides.

The primary structures of the sheep and goat caseinoglycopeptides are very similar as

only residues[1] 113, 141, 144, and 159 are different. The four mutations are conservative and have no influence on the secondary structures. As an example, residues 113 and 159 are located in both proteins in β-turns and present about the same conformational parameters (residue 113, Asn or Asp, P_T = 1.56 or 1.46; f_i = 0.161 or 0.147, respectively; residue 159, Pro or Ser, P_T = 1.52 or 1.43; f_i = 0.102 or 0.120, respectively).

The cow, zebu, and water buffalo caseinoglycopeptides (Mercier et al., 1976) constitute a second group of closely related substances. The primary sequences of the cow and zebu peptides are identical, but between the cow and water buffalo caseinoglycopeptides differences occur at positions 126, 128, 139, 141, 150, 156, 159, 162, and 168. Here again the mutations are conservative when the secondary structures are considered. For example, residue 150 is a β-turn-former in the cow (Pro) or water buffalo (Ser) peptides; in both cases this residue belongs to a β-turn in which the phosphorylated serine residue 151 is located.

When the cow and sheep κ_A-caseinoglycopeptides, representative of each of the two above-mentioned groups, are compared, the number of changes is relatively high; thus it is not surprising that their secondary structures are different. However, despite this fact, the three biologically important parts of these caseinoglycopeptides present the same secondary structure. Indeed, the primary structures of the chymosin-sensitive and sugar-binding sites are the same. The situation is somewhat different when the phosphorylation site is considered, as two changes occur in this area: P.Ser-Pro-Glu-Val (residues 149–152) and P.Ser-Ser-Glu-Ser (residues 151–154) for the cow and sheep peptides, respectively. The mutation corresponding to the $i + 1$ residue of the β-turn decreases the probability of the sheep tetrapeptide adopting a β-turn conformation. An inverse situation is observed when the $i + 3$ residue is considered; thus finally the cow and sheep tetrapeptides adopt a β-turn structure (Table III).

The Insolubility of Para-κ-Casein

The primary structure of para-κ-casein, which constitutes the NH_2-terminal part of κ-casein (Fig. 1), contains a high number of hydrophobic and nonpolar residues and, therefore, its insolubility can be expected. The predicted secondary structure (Fig. 3) is also consistent with the insolubility of para-κ-casein. Indeed, in this part of the casein molecule the presence of an important nucleus of stable antiparallel β-sheets was characterized. More especially, the sequences 22–32 and 40–51 form an antiparallel β-sheet stabilized by the β-turn 35–38. One β-sheet sequence, 73–80, does not form an anti-parallel β-sheet and could therefore play a role in a possible intermolecular association.

The Sequence Surrounding the Chymosin-Sensitive Bond

The specificity of chymosin can be explained in part from the predicted secondary structures of cow and sheep κ_A-caseins. In both proteins the enzyme-sensitive Phe-Met linkage is situated in the sequence 102–108. The latter adopts a helical structure but

[1] The numbering takes into account the position of the caseinoglycopeptide in the whole κ-casein molecule; thus its NH_2-terminal amino acid is residue 106.

TABLE III

COMPARISON BETWEEN THE β-TURNS CORRESPONDING TO THE PHOSPHORYLATION SITES OF COW AND SHEEP κ-CASEINS

β-turn	Residue $i+1$		Residue $i+3$		Values for the β-turns	
	P_T	f_{i+1}	P_T	f_{i+3}	$<P_T>$	$p_t \cdot 10^4$
Cow casein (residues 149–152)						
S P E V	1.52	0.301	0.50	0.053	1.05	1.5
i $i+1$ $i+2$ $i+3$						
Sheep casein (residues 151–154)						
S S E S	1.43	0.139	1.43	0.106	1.28	1.4
i $i+1$ $i+2$ $i+3$						

presents a high $<P_\beta>$ value. If it adopts a β-sheet structure, it should bind the receptor site of chymosin by hydrogen bonds. This conformation should be necessary for the binding to the active site of chymosin, which might also adopt a β-sheet conformation. Moreover, the 102–109 sequence is situated between two stable β-turns, from residues 98–101 with $<P_T> = 1.23, p_t = 2.7.10^{-4}$, and from residues 109–112 with $<P_T> = 1.26, p_t = 2.1.10^{-4}$. It is possible that these β-turns stand out of the casein molecule as a key and thus increase the affinity of chymosin for casein; the existence of a further β-turn, from residues 113 to 116, just adjacent to the latter one, should increase this projection and consequently the affinity of chymosin for casein.

Finally the 98–101 β-turn appears to play a special role: the molecule of cow κ_A-casein possesses only three His residues, situated very close to each other (residues 98, 100, and 102). Two of them (residues 98 and 100) are involved in this 98–101 β-turn and the third (residue 102) belongs to the organized sequence 102–108. The presence of His residues around an active site has been frequently observed and was thus not unexpected. Recently Kaye and Jollès (unpublished results) demonstrated that only one His residue out of the three was essential to observe the full clottability of cow κ_A-casein by chymosin. The use of diethyl pyrocarbonate indicated that two His residues were modified at low reagent concentrations. It is possible to suggest that they are situated in the β-turn and may therefore be more readily reactive. The third His residue, which reacted only at high reagent concentrations and seemed essential for the chymosin action, may be situated in the organized sequence. In sheep κ_A-casein a fourth His residue was characterized (residue 133), but did not seem to play a role either in biological activity or in the ordering of the secondary structure.

The authors thank Mrs. M. P. Hildebrand for skillful technical assistance.

This research was supported by the Institut National de la Santé et de la Recherche Médicale (Unité 124 and group U-116) and the Centre National de la Recherche Scientifique (E.R. 102).

Received for publication 28 December 1977 and in revised form 20 March 1978.

REFERENCES

ALAIS, C., and P. JOLLÈS. 1961. Etude comparée des caséino-glycopeptides formés par action de la présure sur les caséines de vache, de brebis et de chèvre. II. Etude de la partie non-peptidique. *Biochim. Biophys. Acta.* **51**:315.

ALAIS, C., and P. JOLLÈS. 1967. Isolation, purification and analysis of two κ-casein-like fractions from sheep casein. *J. Dairy Sci.* **50**:1555.

ARGOS, P., J. SCHWARZ, and J. SCHWARZ. 1976. An assessment of protein secondary structure prediction methods based on amino acid sequence. *Biochim. Biophys. Acta.* **439**:261.

AUBERT, J. P., and M. H. LOUCHEUX-LEFEBVRE. 1976. Conformational study of α_1-acid glycoprotein. *Arch. Biochem. Biophys.* **175**:400.

AUBERT, J. P., G. BISERTE, and M. H. LOUCHEUX-LEFEBVRE. 1976. Carbohydrate-peptide linkage in glycoproteins. *Arch. Biochem. Biophys.* **175**:410.

BEELEY, J. G. 1976. Location of the carbohydrate groups of ovomucoid. *Biochem. J.* **159**:335.

BEELEY, J. G. 1977. Peptide chain conformation and the glycosylation of glycoproteins. *Biochem. Biophys. Res. Commun.* **76**:1051.

BRIGNON, G., J.-C. MERCIER, B. RIBADEAU-DUMAS, and B. C. DAS. 1972. Structure primaire de la paracaséine κ bovine. *(Fed. Eur. Biochem. Soc.) FEBS Lett.* **27**:301.

CHEN, Y. H., J. T. YANG, and K. H. CHAU. 1974. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry.* **13**:3350.

CHOU, P. Y., and G. D. FASMAN. 1974. Conformational parameters for amino acid in helical, β-sheet, and random coil regions calculated from proteins. *Biochemistry.* **13**:211.

CHOU, P. Y., and G. D. FASMAN. 1977. β-Turns in proteins. *J. Mol. Biol.* **115**:135.

CHOU, P. Y., A. J. ADLER, and G. D. FASMAN. 1975. Conformational prediction and circular dichroism studies on the *lac* repressor. *J. Mol. Biol.* **96**:29.

DAYHOFF, M. O. 1976. Atlas of protein sequence and structure. National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C. **5**:suppl. 2.

DELFOUR, A., J. JOLLÈS, C. ALAIS, and P. JOLLÈS. 1965. Caseino-glycopeptides: characterization of a methionine residue and of the N-terminal sequence. *Biochem. Biophys. Res. Commun.* **19**:452.

FASMAN, G. D., P. Y. CHOU, and A. J. ADLER. 1976. Prediction of the conformation of the histones. *Biophys. J.* **16**:1201.

JOLLÈS, P. 1966. Progress in the chemistry of casein. *Angew. Chem. Int. Ed. Engl.* **5**:558.

JOLLÈS, P. 1975. Structural aspects of the milk clotting process. Comparative features with the blood clotting process. *Mol. Cell. Biochem.* **7**:73.

JOLLÈS, J., F. SCHOENTGEN, C. ALAIS, and P. JOLLÈS. 1972. Studies on the primary structure of cow κ-casein: the primary sequence of para-κ-casein. *Chimia.* **26**:645.

JOLLÈS, J., A.-M. FIAT, C. ALAIS, and P. JOLLÈS. 1973. Comparative study of cow and sheep κ-caseinoglycopeptides: determination of the N-terminal sequences with a Sequencer and location of the sugars. *Fed. Eur. Biochem. Soc. (FEBS) Lett.* **30**:173.

JOLLÈS, J., A.-M. FIAT, F. SCHOENTGEN, C. ALAIS and P. JOLLÈS. 1974. The amino acid sequence of sheep κ_A-casein. II. Sequence studies concerning the κ_A-caseinoglycopeptide and establishment of the complete primary structure of the protein. *Biochim. Biophys. Acta* **365**:335.

LEWIS, P. N., F. A. MOMAMY, and H. A. SCHERAGA. 1973. Chain reversals in proteins. *Biochim. Biophys. Acta.* **303**:211.

LOWRY, O. H., N. J. ROSEBROUGH, A. L. FARR, and R. J. RANDALL. 1951. Protein determination with the Folin phenol reagent. *J. Biol. Chem.* **193**:265.

MCKENZIE, H. A., and R. G. WAKE. 1961. An improved method for the isolation of κ-casein. *Biochim. Biophys. Acta.* **47**:240.

MERCIER, J. C., J. URO, B. RIBADEAU-DUMAS, and F. GROSCLAUDE. 1972. Structure primaire du caséinomacropeptide de la caséine κB_1 bovine. *Eur. J. Biochem.* **27**:535.

MERCIER, J. C., J. M. CHOBERT, and F. ADDEO. 1976. Comparative study of the amino acid sequences of the caseinomacropeptides from seven species. *Fed. Eur. Biochem. Soc. (FEBS) Lett.* **72**:208.

NAGARAJAN, M., and V. S. R. RAO. 1977. Conformational analysis of glycoproteins. Part I. Conformation of the protein segment at the site of peptide-sugar linkage. *Curr. Sci.* **46**:395.

SMALL, D., P. Y. CHOU, and G. D. FASMAN. 1977. Occurrence of phosphorylated residues in predicted β-turns: Implications for β-turn participation in control mechanisms. *Biochem. Biophys. Res. Commun.* **79**:341.

THIERY, J.-M. 1969. Thesis, University of California at Berkeley. 69–18978. University Microfilms, Ann Arbor, Mich.