Bioinformation
www.bioinformation.net

open access
Hypothesis

Analysis of casein alpha S1 & S2 proteins from different mammalian species

Tariq Ahmad Masoodi*, Gowhar Shafi

Institute of Genetics and Hospital for Genetic Diseases, Begumpet, Hyderabad 500016, India; Tariq Ahmad Masoodi – Email: tahamasoodi@gmail.com; *Corresponding author

Received October 04, 2009; accepted November 15, 2009; published March 31, 2010

Abstract

Nowadays, the quality of any food used for human consumption is, to a considerable extent, considered by its possible contribution to the maintenance or improvement of the consumer's health. In developed countries there is increasing interest in goat milk and its derivates, the quality of which is considered of special importance in the light of current tendencies favouring healthy eating. In particular, goat's milk is a hypoallergenic alternative to cow's milk in the human diet. In the present work, we studied the casein alpha S1 and S2 proteins of cow, goat and sheep for comparative analysis. We found that the amino acid sequence of these proteins is almost same in goat and sheep but there are several changes at different base pairs when these two sequences are compared with cow. Prediction of secondary structures (GOR) was performed for alpha s1 and s2 proteins to gain functional insights. Our in silico study revealed considerable identity in chemical properties between goat and sheep but a considerable dissimilarity in cow with goat and sheep casein proteins. Moreover, the effect amino acid change on secondary structures in the three species is discussed. There was no significant difference found between goat and sheep alpha S1 and S2 proteins, so naturally both will be having same properties. The study concludes that sheep milk is another convenient alternative for the cow milk allergic children.

Keywords: casein, alpha proteins, sequence analysis,

Background:

The best nutritional option for newborn infants is mother's milk; however some infants may not be exclusively breast fed during the first months of life. In that case, another substitute or alternative must be provided as cow milk. This substitution results in an allergic disease known as cow milk protein allergy (CMPA) in 2–6% of children **[1].** Nowadays, most common alternatives are soy and extensively hydrolyzed milk proteins formulae **[1].** However, there is evidence that 10–20% of children allergic to cow milk do not tolerate soy derivatives **[2-4]** and some cases of high immunological reaction to extensively hydrolyzed formulae have been reported **[5-7]**. Immunoglobulin E-mediated allergies belong to the most common forms of immunologically mediated forms of hypersensitivity reactions to food **[8].** In sensitized individuals dietary intake of food can cause a variety of clinical manifestations reaching from oral allergy syndrome and gastrointestinal symptoms (e.g., vomiting, diarrhoea) to skin, respiratory, and severe systemic manifestations such as anaphylactic shock **[9– 11]**. The development of food allergy shows a typical course **[12].** It starts early in childhood mainly against Ags encountered in the initial diet (e.g., cow's milk, eggs) and affects between 4% and 6% of children. Milk is one of the first food components introduced into the diet and therefore represents one of the most important food allergen sources in terms of frequency and severity of allergic manifestations **[13 –16].** The symptoms of cow's milk allergy are due to IgE-mediated activation of mast cells and basophils as well as activation of allergen-specific T cells, and they comprise a plethora of gastrointestinal, skin, respiratory, and severe systemic manifestations such as death due to anaphylactic shock. Unlike in respiratory allergy, which proceeds untreated from mild (e.g., rhinoconjunctivitis) to severe manifestations (e.g., asthma), many milk-allergic children grow out their allergy, and induction of tolerance against cow's milk allergens has already been described **[17, 18].** Cow's milk contain more than 25 different proteins, but only the whey proteins alpha-lactalbumin, beta-lactoglobulin, BSA, and lactoferrin, as well as the four caseins, have been identified as allergens **(19).** The casein fraction is composed of alpha S1-, alpha S2-, beta-, and k-casein, of which alpha S1-casein seems to be a major allergen according to IgE and T cell recognition data **[20–23].** In developed countries there is increasing interest in goat milk and its derivates, the quality of which is considered of special importance in the light of current tendencies favouring healthy eating. In particular, the composition of goat milk is said to have certain advantages over that of cow milk, and thus the former is preferable for some consumers **[24-26].** On these bases, the identification of a suitable protein source for children allergic to cow milk represents an important goal for both nutritionists and paediatricians. Therefore, the present study aims at evaluating the suitability of sheep milk for the nutrition of children allergic to cow milk by analysing primary structure, secondary structure and hydropathicity plots of sheep, goat and cow milk proteins.

Methodology:

Dataset

Casein alpha S1 and S2 proteins from sheep were analyzed in the present study. The protein sequences are available at NCBI with accession numbers [GenBank: P04653 and CAA26983 respectively]. The casein alpha S1 and S2 protein sequences of goat and cow were downloaded from NCBI [http://www.ncbi.nlm.nih.gov] and summary of sequence data is given in **Table 1 (see supplementary material)**. The dataset consists of casein alpha S1 and S2 protein sequences from sheep, goat and cow.

Physico-chemical analyses:

The ProtParam tool at ExPASy [http://www.expasy.ch/] was used to analyse amino acid and atomic compositions, isoelectric point, extinction coefficient and hydropathicity in the three species.

Multiple Sequence Alignment & Secondary structure prediction

The CLUSTALW at European Bioinformatics Institute[http://www.ebi.ac.uk/Tools/clustalw] was used for generating a multiple sequence alignment [MSA] of casein alpha S1 and S2 protein sequences from three different species and the GOR at EXPASY [http://npsa-pbil.ibcp.fr] to assign secondary structures to casein alpha S1 and S2 sequences.

Results:

Physico-chemical analysis

The casein alpha S1 protein consists of 214 amino acids in each species while casein alpha S2 consists of 223 residues in sheep and goat and 222 in cow. Besides the identical sequence alignment between the casein alpha S1 and S2 proteins of sheep and goat, these proteins also share considerable equality in their percentage content of amino acids as given in **Table 2 and 3 (see supplementary material)**.

```
MKLLILTCLVAVALARPKHPINHRGLSPEVPNENLLRFVVAPFPEVFRKE 50 gi|416750|sp|P18626
MKLLILTCLVAVALARPKHPIKHQGLSSEVLNENLLRFVVAPFPEVFRKE 50 gi|57015287|sp|P04653
MKLLILTCLVAVALARPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKE 50 gi|115646|sp|P02662
*********************:*:**. ** *******.******** **

NINELSKDIGSESTEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKED 100 gi|416750|sp|P18626
NINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKED 100 gi|57015287|sp|P04653
KVNELSKDIGSESTEDQAMEDIKQMEAESISSSEEIVPNSVEQKHIQKED 100 gi|115646|sp|P02662
::*********** ******* ***:* * **********.***:*****

VPSERYLGYLEQLLRLKKYNVPQLEIVPKSAEEQLHSMKEGNPAHQKQPM 150 gi|416750|sp|P18626
VPSERYLGYLEQLLRLKKYNVPQLEIVPKSAEEQLHSMKEGNPAHQKQPM 150 gi|57015287|sp|P04653
VPSERYLGYLEQLLRLKKYKVPQLEIVPNSAEERLHSMKEGIHAQQKEPM 150 gi|115646|sp|P02662
*******************:********:****:*******  *:**:**

IAVNQELAYFYPQLFRQFYQLDAYPSGAWYYLPLGTQYTDAPSFSDIPNP 200 gi|416750|sp|P18626
IAVNQELAYFYPQLFRQFYQLDAYPSGAWYYLPLGTQYTDAPSFSDIPNP 200 gi|57015287|sp|P04653
IGVNQELAYFYPELFRQFYQLDAYPSGAWYYVPLGTQYTDAPSFSDIPNP 200 gi|115646|sp|P02662
*.**********:******************:******************

IGSENSGKTTMPLW 214 gi|416750|sp|P18626
IGSENSGKITMPLW 214 gi|57015287|sp|P04653
IGSENSEKTTMPLW 214 gi|115646|sp|P02662
****** * *****
```

Figure 1: Multiple sequence alignment of casein alpha S1 in the three species

```
MKFFIFTCLLAVALAKHKMEHVSSSEEPINIFQEIYKQEKNMAIHPRKEK 50 gi|15706248|emb|CAC21704
MKFFIFTCLLAVALAKHKMEHVSSSEEPINISQEIYKQEKNMAIHPRKEK 50 gi|732894|emb|CAA26983
MKFFIFTCLLAVALAKNTMEHVSSSEESI-ISQETYKQEKNMAINPSKEN 49 gi|27806963|ref|NP_776953
****************:.*********.* * ** *********:* **:

LCTTSCEEVVRNANEEEYSIRSSSEESAEVAPEEIKITVDDKHYQKALNE 100 gi|15706248|emb|CAC21704
LCTTSCEEVVRNADEEEYSIRSSSEESAEVAPEEVKITVDDKHYQKALNE 100 gi|732894|emb|CAA26983
LCSTFCKEVVRNANEEEYSIGSSSEESAEVATEEVKITVDDKHYQKALNE 99  gi|27806963|ref|NP_776953
**:* *:******:****** **********.**:***************

INQFYQKFPQYLQYPYQGPIVLNPWDQVKRNAGPFTPTVNREQLSTSEEN 150 gi|15706248|emb|CAC21704
INQFYQKFPQYLQYLYQGPIVLNPWDQVKRNAGPFTPTVNREQLSTSEEN 150 gi|732894|emb|CAA26983
INQFYQKFPQYLQYLYQGPIVLNPWDQVKRNAVPITPTLNREQLSTSEEN 149 gi|27806963|ref|NP_776953
************** ***************** *:***:***********

SKKTIDMESTKVFTKKTKLTEEEKNRLNFLKIISQYYQKFAWPQYLKTVD 200 gi|15706248|emb|CAC21704
SKKTIDMESTEVFTKKTKLTEEEKNRLNFLKKISQYYQKFAWPQYLKTVD 200 gi|732894|emb|CAA26983
SKKTVDMESTEVFTKKTKLTEEEKNRLNFLKKISQRYQKFALPQYLKTVY 199 gi|27806963|ref|NP_776953
****:*****:******************** *** ***** *******

QHQKAMKRWTQPKTNAIPYVRYL 223 gi|15706248|emb|CAC21704
QHQKAMKPWTQPKTNAIPYVRYL 223 gi|732894|emb|CAA26983
QHQKAMKPWIQPKTKVIPYVRYL 222 gi|27806963|ref|NP_776953
******* * ****:.*******
```

Figure 2: Multiple sequence alignment of casein alpha S2 in the three species

Sequence analyses and secondary structure prediction

The alpha S1 and S2 casein protein sequences of cow were compared with S1 and S2 of goat and sheep as given in Table 4 & 5 respectively. The alpha S1 and S2 protein sequences of goat and sheep obtained from GenBank have at least 99% similarity among themselves but the two differ markedly from the cow. A multiple sequence alignment is obtained using the software ClustalW **(Figure 1).** Changes in sequences between the cow and other two species are summarized in **Table 2 (see supplementary material)** with changes in corresponding predicted secondary structures. On comparing sequences of S1 and S2 proteins by using the program ClustalW, it was observed that there are almost no changes in the sequences and the secondary structure between goat and cow. However when the two were compared with cow, it was found there is a dramatic change in the cow protein sequences with a definite change in the secondary structure and the chemical properties as shown in table 4 and 5 for S1 and S2 respectively.

Discussion:

Cow's milk contains more than 25 different proteins **[19]** but the casein fraction composed of alpha S1-, alpha S2-, beta-, and k-casein particularly alpha S1-casein seems to be a major allergen **[20 – 23].** The chemical property and biological function of a protein is a direct consequence of its primary structure **[27].** The most common and generally more accessible approach to protein function prediction is 'inheritance through homology' - that is, the knowledge that proteins with similar sequences frequently carry out similar functions **[28].** For a majority of proteins it is already possible to predict their approximate function with reasonable accuracy based on their evolutionary relationship or sequence similarity to proteins with known functions **[29-31].**

Since the casein alpha S1 and S2 from goat and sheep share a great similarity so they could be thought of having the same effect and function. Much anecdotal evidence is available, suggesting that goat's milk is a hypoallergenic alternative to cow's milk in the human diet. Improvement in the symptoms associated with colic and minor digestive disorders, asthma and eczema have all been reported **[32].** Scientific and clinical studies also suggest that infants and children who are sensitive to a cow's milk based product often thrive better when goat milk based product is substituted **[27].**

Goat milk as a substitute for cow milk was studied in 38 children during a 5 months period **[33].** The children on goat milk surpassed those on cow milk in weight gain, height, skeletal mineralization, and blood serum contents of Vitamin A, calcium, thiamin, riboflavin, niacin and hemoglobin. Similar findings were obtained in studies with rats **[34].** In French clinical studies over 20 years with cow milk allergy patients the conclusion was that substitution with goat milk was followed by "undeniable" improvements **[33].** In other French extensive clinical studies with children allergic to cow milk, the treatment with goat milk produced positive results in 93% of the children and was recommended as a valuable aid in child nutrition because of less allergenicity and better digestibility than cow milk **[33].**

Conclusion:
Since casein alpha S1 and S2 proteins of sheep are almost identical to goat, so sheep milk can be a useful another convenient alternative to the cow milk allergic children.

References:
[1] EI El-Agamy, Small Rum Res (2007) 68: 64-72
[2] L Businco et al., J. Pediatr (1992) 121: S21-S28 [PMID:1447630]
[3] J Maldonado et al., Early Hum Dev (1998) 53 S23-S32 [PMID: 10102652]
[4] RS Zeiger et al., J Pediatr (1999) 134: 614-622 [PMID: 10228298]
[5] L Businco et al., Ann Allergy (1989) 62: 333-335 [PMID: 2705659]
[6] HA Sampson et al., Pediatrics (1992) 90: 463-465 [PMID: 1518709]
[7] D de Boissieu & C Dupont, J Pediatr (2002) 141: 271-273 [PMID: 12183726]
[8] S Bischoff & SE Crowe, Gastroenterology (2005) 128: 1089-1113 [PMID: 15825090]
[9] PD Buisseret, Lancet (1978) 1: 304-305 [PMID: 75338]
[10] T Hofer & B Wuthrich, Schweiz Med Wochenschr (1985) 115: 1437-1442 [PMID: 4071016]
[11] SA Bock et al., J Allergy Clin Immunol (2001) 107: 191-193 [PMID: 11150011]
[12] MR Kulig et al., J Allergy Clin Immunol (1999) 103: 1173-1179 [PMID: 10359902]
[13] MH Loveless, J Allergy (1950) 21: 489-499 [PMID: 14794310]
[14] W Burks BK Ballmer-Weber, Mol Nutr Food Res (2006) 50: 595-603 [PMID: 16810732]
[15] SH Sicherer & HA Sampson, J Allergy Clin Iimmunol (2006) 117: 470-475 [PMID: 20042231]
[16] A Finkelstein, Kinderheilkunde (1905) 4: 64-72
[17] G Finzio, Pediatria (1911) 19: 641
[18] M Morisset et al., Eur Ann Allergy Clin Immunol (2007) 39: 12-19 [PMID: 17375736]
[19] JM Wal, Allergy (1998) 53: 1013-1022
[20] P Spuergin et al., Allergy (1996) 51: 306-312 [PMID: 8836334]
[21] P Chatchatee et al., J Allergy Clin Immunol (2001) 107: 379-383 [PMID: 11174208]
[22] B Ruiter et al., Clin Exp Allergy (2006) 36: 303-310 [PMID: 16499640]
[23] B Ruiter et al., Int Arch Allergy Immunol (2007) 143: 119-126 [PMID: 17228168]
[24] MR Sanz Sampelayo et al., Small Ruminant Research (2002) 43: 141-148
[25] MJM Alférez et al., J Dairy Res (2001) 64: 451-461[PMID: 11694047]
[26] M Barrionuevo et al., J Dairy Sci (2002) 85: 657-664 [PMID: 11949871]
[27] K Gupta et al., BMC Bioinformatics (2005) 6:105 [PMID: 15850477]
[28] D Lee et al., Nat Rev Mol Cell Biol (2007) 8:995-1005 [PMID: 18037900]
[29] JC Whisstock & AM Lesk, Q Rev Biophys (2003), 36: 307-340 [PMID: 15029827]
[30] PD Dobson et al., Curr Med Chem (2004) 11: 2135-2142 [PMID: 15279553]
[31] T Doerks et al., Nucleic Acids Res (2004) 32:6321-6326 [PMID: 15576358]
[32] YW Park, Small Ruminant Research (1994) 14: 151-159
[33] GFW Haenlein, Small Ruminant Research (2004) 51: 155-163
[34] YW Park et al., J Dairy Sci (1986) 69 2608-2615 [PMID: 3805444]

Edited by P. Kangueane
Citation: Masoodi *et al,* Bioinformation 4(9): 430-435 (2010)

Supplementary material:

Table 1: Selected proteins for comparison analysis

Genbank Accession	Protein Name	Source
P02662	Casein alpha S1	Cow
P18626	Casein alpha S1	Goat
P04653	Casein alpha S1	Sheep
NP_776953	Casein alpha S2	Cow
CAC21704	Casein alpha S2	Goat
CAA26983	Casein alpha S2	Sheep

Table 2: ProtScale result for amino acid composition of alpha S1 protein

	Cow	Goat	Sheep
Ala (A)	5.6%	7.0%	7.0%
Arg (R)	2.8%	3.3%	2.8%
Asn (N)	3.7%	5.1%	4.7%
Asp (D)	3.3%	3.3%	3.3%
Cys (C)	0.5%	0.5%	0.5%
Gln (Q)	6.5%	6.5%	7.0%
Glu (E)	11.7%	9.3%	9.3%
Gly (G)	4.2%	4.2%	4.2%
His (H)	2.3%	1.9%	1.9%
Ile (I)	5.6%	5.6%	4.7%
Leu (L)	10.3%	10.3%	10.7%
Lys (K)	7.0%	6.5%	7.0%
Met (M)	2.8%	2.8%	2.8%
Phe (F)	3.7%	3.3%	3.3%
Pro (P)	7.9%	8.9%	7.9%
Ser (S)	7.5%	8.4%	8.9%
Thr (T)	2.8%	2.8%	1.9%
Trp (W)	0.9%	0.9%	0.9%
Tyr (Y)	4.7%	5.1%	5.1%
Val (V)	6.1%	5.1%	5.1%

Table 3: Amino acid composition of casein alpha S2 protein

	Cow	**Goat**	**Sheep**
Ala (A)	5.0%	5.4%	5.4%
Arg (R)	2.7%	3.6%	3.1%
Asn (N)	6.3%	5.8%	5.4%
Asp (D)	1.8%	2.2%	2.7%
Cys (C)	1.4%	1.3%	1.3%
Gln (Q)	7.2%	7.2%	7.2%
Glu (E)	10.8%	10.8%	11.2%
Gly (G)	0.9%	0.9%	0.9%
His (H)	1.4%	2.2%	2.2%
Ile (I)	5.4%	6.3%	5.4%
Leu (L)	7.2%	5.8%	6.3%
Lys (K)	11.3%	11.2%	11.2%
Met (M)	2.3%	2.2%	2.2%
Phe (F)	4.1%	4.5%	4.0%
Pro (P)	4.5%	5.4%	5.4%
Ser (S)	7.7%	6.3%	6.7%
Thr (T)	7.2%	6.7%	6.7%
Trp (W)	0.9%	1.3%	1.3%
Tyr (Y)	5.4%	5.4%	5.4%
Val (V)	6.8%	5.4%	5.8%

Table 4: Comparison of amino acid residues in casein alpha S1 protein of cow with other two species giving its specific position with changes in secondary structure and their properties

Base change	**Change in aa of alpha s2 in cow**	**Change in properties**	**Change in secondary structure**
Alpha s2 of goat			
17	N→H	Hydrophilic	C→H
18	T→K	Hydrophilic	H
28	S→P	Hydrophilic	H→C

32	S→F	Hydrophilic→Hydrophobic	H
35	T→I	Hydrophilic→Hydrophobic	H
45	N→H	Hydrophilic	C→H
47	S→R	Hydrophilic	C
50	N→K	Hydrophilic	C→E
55	F→S	Hydrophobic→Hydrophilic	C
57	K→E	Hydrophilic	C
71	G→R	Hydrophilic	C→H
82	T→P	Hydrophilic	H
85	V→I	Hydrophobic	H
115	L→P	Hydrophobic→Hydrophilic	C
133	V→G	Hydrophobic→Hydrophilic	C
135	I→F	Hydrophobic	C
139	L→V	Hydrophobic	C
155	V→I	Hydrophobic	E
161	E→K	Hydrophilic	H
182	K→I	Hydrophilic→Hydrophobic	H
186	R→Y	Hydrophilic	H→C
192	L→W	Hydrophobic→Hydrophilic	C
200	Y→D	Hydrophilic	H
208	P→R	Hydrophilic	C→H
210	I→T	Hydrophobic→Hydrophilic	C→H
215	K→N	Hydrophilic	C
216	V→A	Hydrophobic	E→C
Alpha s2 of sheep			
17	N→H	Hydrophilic	C→H
18	T→K	Hydrophilic	H
28	S→P	Hydrophilic	H→C
35	T→I	Hydrophobic→Hydrophilic	H
45	N→H	Hydrophilic	C→H
47	S→R	Hydrophobic	C
50	N→K	Hydrophobic	C→E
55	F→S	Hydrophobic	C
57	K→E	Hydrophilic	C
64	N→D		C
71	G→R	Hydrophilic	C→H
82	T→P	Hydrophobic→Hydrophilic	H
133	V→G	Hydrophilic	C
135	I→F	Hydrophobic	C
139	L→V	Hydrophobic	C
155	V→I	Hydrophilic	E
186	R→Y	Hydrophilic	H→E
192	L→W	Hydrophilic	C
200	Y→D	Hydrophilic→Hydrophobic	H
210	I→T	Hydrophobic	C
215	K→N	Hydrophilic	C
216	V→A	Hydrophobic	E→C

Table 5: Comparison of amino acid residues in casein alpha S1 protein of cow with other two species giving its specific position with changes in secondary structure and their properties

Base change	Change in aas of alpha s1			Change in property	Change in secondary structure
	cow		goat		
22	K	→	N	Hydrophilic	C
24	Q	→	R	Hydrophilic	C
27	P	→	S	Hydrophilic	C
28	Q	→	P	Hydrophilic	H→C
31	L	→	P	Hydrophobic→Hydrophilic	H→C
39	F	→	V	Hydrophobic	C→H
48	G	→	R	Hydrophilic	C→H
51	K	→	N	Hydrophilic	H→C
52	V	→	I	Hydrophobic	H
72	I	→	A	Hydrophobic	C
76	E	→	K	Hydrophilic	H
78	E	→	G	Hydrophilic	H→C
80	I	→	S	Hydrophobic→Hydrophilic	C
91	V	→	A	Hydrophobic	H
95	H	→	Y	Hydrophobic→Hydrophilic	H

120	K	→	N	Hydrophobic	C
129	N	→	K	Hydrophilic	C
134	R	→	Q	Hydrophilic	H
142	I	→	N	Hydrophobic→Hydrophilic	H→C
143	H		P	Hydrophilic	H→C
145	Q	→	H	Hydrophilic	H→C
148	E	→	Q	Hydrophilic	C
152	G	→	A	Hydrophilic→Hydrophobic	E→H
163	E	→	Q	Hydrophilic	H
182	V	→	L	Hydrophobic	E
207	E	→	G	Hydrophilic	E→C
	cow		sheep		
28	Q	→	S	Hydrophilic	H
39	F	→	V	Hydrophobic	C→H
48	G	→	R	Hydrophilic	C→H
51	K	→	N	Hydrophilic	H→C
52	V	→	I	Hydrophobic	H
64	T		I	Hydrophilic→Hydrophobic	C→H
72	I	→	A	Hydrophobic	H
76	E	→	K	Hydrophilic	H
78	E	→	G	Hydrophilic	H→C
80	I	→	S	Hydrophobic→Hydrophilic	C
91	V	→	A	Hydrophobic	H
95	H	→	Y	Hydrophilic	H
120	K	→	N	Hydrophilic	C
129	N	→	K	Hydrophilic	C
134	R	→	Q	Hydrophilic	H
142	I	→	N	Hydrophobic→Hydrophilic	H→C
143	H		P	Hydrophilic	H→C
145	Q	→	H	Hydrophilic	H→C
148	E	→	Q	Hydrophilic	C
152	G	→	A	Hydrophilic→Hydrophobic	E→H
163	E	→	Q	Hydrophilic	H
182	V	→	L	Hydrophobic	E
207	E	→	G	Hydrophilic	E
209	T		I	Hydrophilic→Hydrophobic	E