Probing β-casein structure using Nuclear Magnetic Resonance (NMR) and Fourier Transform Infrared (FTIR) spectroscopy

Tatijana Markoska

Thesis submitted in fulfilment of the requirements for the degree of

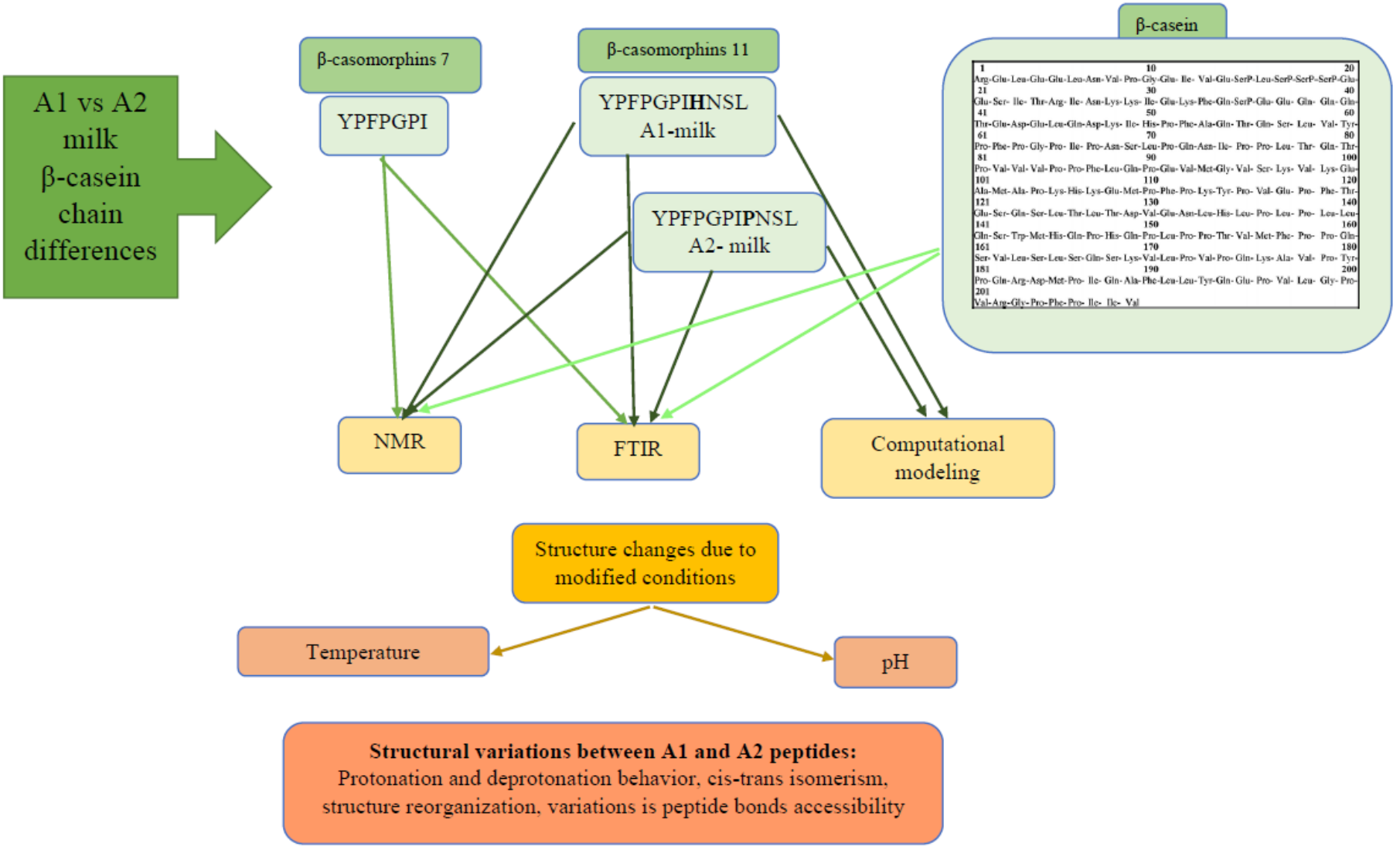
Doctor of Philosophy

Victoria University, Australia

Institute of Sustainable Industries and Liveable Cities

March 2023

CONCEPTUAL FRAMEWORK OF THE STUDY



ABSTRACT

Milk proteins are essential nutritive components and have great impact on the quality of dairy processing. Milk proteins are classified into two groups or caseins and whey proteins. Among the caseins, the second most abundant casein is β-casein. In the recent years, the difference in the genetic variants of β-casein and its impact on human health and milk processing has led to an increased interest among researchers. Thus, individual genetic variants have shown dissimilarities in the physicochemical behaviour during processing. The commercially most common β-casein milk variants are A1 and A2 milk. The difference between these variants is in the amino acid position 67 in the β-casein polypeptide chain with histidine in A1 milk and proline in A2 milk. Concerning the milk processing, A1 casein have shown to have better coagulation characteristics compared to A2 milk which is important for production of fermented products like cheese or yoghurt. However, in regards to the human consumption epidemiological evidence debate that consumption of A1 β-casein may be a risk factor for many diseases although the literature does not support these claims. Moreover, this appears to be due to the proteolysis of β-casein into smaller peptides in the human gut or bioactive peptides known as β-casomorphins (BCM).

The most common peptides are BCM5, BCM7, BCM11 and BCM15. The most studied peptide to date is BCM7 that have been characterised to have opioid activity by binding to the μ-receptors in the human body and leads to a range of physiological responses. However, the structure of the peptide in conditions that have importance for its cleavage in human gut or during milk processing is lacking in the literature. In the current study, we observed that BCM7 shown structural difference under modified pH conditions including pH 2.3 and pH 7.0 that are native environmental condition of human gut and thus gives maximum activity for enzymatic cleavage in the digestive track. These structural differences were observed to exist due to the *cis-trans* isomerisam of the X-Pro bond. The finding shown pH dependence of the peptides bond accessibility for cleavage into smaller particles.

The most important peptide concerning the difference between A1 and A2 β casein is BCM11. BCM11 from A1 milk have histidine in position 8 of the polypeptide chain and BCM11 from A2 milk have proline in the same position of the polypeptide chain. These amino acids are the major factor contributing to the different properties of β casein molecule originating from A1 or A2 milk.

By observing the structure, we verified that the peptides adapt different conformers that further rearrange when pH and temperature conditions are modified. This was complemented with higher structural flexibility for BCM11 peptide from A1 β-casein due to loose conformation and tautomeric behaviour of imidazole ring of histidine that was dependent on the applied conditions. These structural differences have great effect on the accessibility of peptide bonds for cleavage into smaller peptides. Further computational studies of these peptides confirmed that the imidazole ring of the histidine have important role on the structural stability of the molecule involving changed charge transfer and affecting the acid-base chemistry of the molecule. In addition, the accessibility of the peptide bond for proteolysis can be dependent on the structural orientation of the β-casein. The proline in position 67 in A2 β-casein can be the factor for low accessibility of the bond 66-67 for cleavage. This is due to greater involvement of proline in turns formation in molecules. Moreover, the structural conformation of the molecules are also dependent of the environment. The predominant factor is the pH since the enzymes have optimal activity on defined pH values. Thus, the structural conformation of the β-casomorphin peptides can be pH dependent.

The effect of the temperature and pH on the β-casein molecule was observed to have large effect on the secondary structure of the molecule. The findings confirmed loss of well defined structural components (α-helix) and formation of random structures when pH was increased to neutral and temperature lowered to 4 ^{o}C. The modified conditions that are essential for milk processing disturbed the native hydrogen bonds and promoted formation of new structures. The changed conditions were also observed to have different effect on the β-casein genotype. β-Casein from A1 and A2 milk adapt different conformational states. Hence, β-casein from A2 have open and porous protein network with native aggregated β-sheets, however, β-casein from A1 milk adapts more dense structure with small network pores. The effect of pH largely affects the conformational orientation of both genotypes with formation of large irregular aggregates in A2 β-casein and more dense aggregated β-sheets in A1 β-casein. Observing the structural features of the peptides and β-casein can assist in understanding the reasons for their specific behaviour during milk processing or human digestion.

DECLARATION

"I, Tatijana Markoska, declare that the PhD thesis entitled "Probing casein structure using Nuclear magnetic resonance (NMR) and Fourier transform infrared (FTIR) spectroscopy" is no more than 80,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

"I have conducted my research in alignment with the Australian Code for the Responsible Conduct of Research and Victoria University's Higher Degree by Research Policy and Procedures.

Signature

Date

30/03/2023

ACKNOWLEDGEMENTS

First and foremost, I would like to express my profoundest gratitude to my principal supervisor, Professor Todor Vasiljevic, for his professional guidance, motivation, patience, immeasurable knowledge and encouragement throughout the course of my study. In the past few years his guidance and supervision helped me to become confident and capable to work independently on my research. I am proud and honoured to be supervised by an academic of his calibre. Thank you for all opportunities and understanding.

My gratitude and appreciation also goes to my associate supervisor, Professor Thom Huppertz for his academic support, supportive guidance, motivating discussions that expanded my professional development. I am also grateful to Professor Andrew Smallridge for his scientific, technical support and for guiding me throughout NMR instrumentation. I gratefully acknowledge the scientific contribution and professional guidance by Professor John Orbell.

I sincerely appreciate the technical support by the technical staff in Werribee and Footscray Park campuses. The gratitude goes to Stacey Lloyd, Melinda Tagiwasa, Chairmane DiQuattro, Mary Marshall, Nishantha Illangantilaka and Larruceo Bautista. Special thanks to Muditha Dissanayake for always being supportive and for training me and assisting me in developing and learning different scientific protocols.

I would like to express my profundes gratitude to my fellow colleagues, Davor Daniloski, Elaheh Ahmadi, Anushka Mediwaththe, Geetha Vermula, Surjit Kaur, Derrick D'Souza, Dimuthu Bogahawaththa and all others for their support and friendship.

Finally, I would like to express my most profound regards to my sons, Borjan and Martin and my husband Bojan for being supportive and understanding throughout this research.

VICTORIA UNIVERSITY

DETAILS OF INCLUDED PAPERS: THESIS WITH PUBLICATION

Please list details of each scholarly publication and/or manuscript included in the thesis submission. Copies of published scholarly publications and/or manuscripts submitted and/or final draft manuscripts should also be included in the thesis submission.

This table must be incorporated in the thesis before the Table of Contents.

Chapter No.	Publication Title	Publication Status • Published • Accepted for publication • In revised and resubmit stage • Under review • Manuscript ready for submission	Publication Details • Citation, if published • Title, Journal, Date of acceptance letter and Corresponding editor's email address • Title, Journal, Date of submission
2B	Unravelling Conformational Aspects of Milk Protein Structure—Contributions from Nuclear Magnetic Resonance Studies.	Published	Markoska, T., Vasiljevic, T. and Huppertz, T., 2020. Unravelling Conformational Aspects of Milk Protein Structure—Contributions from Nuclear Magnetic Resonance Studies. Foods, 9(8), p.1128.
3	pH-induced changes in β-casomorphin 7 structure studied by 1H nuclear magnetic resonance and Fourier-transform infrared spectroscopy.	Published	Markoska, T., Huppertz, T. and Vasiljevic, T., 2021. pH-induced changes in β-casomorphin 7 structure studied by 1H nuclear magnetic resonance and Fourier-transform infrared spectroscopy. International Dairy Journal, 121, p.105106.
4	Temperature- and pH-induced structural changes of β-casomorphin 11-A1 and -A2 studied by nuclear magnetic resonance, Fourier-transform infrared spectroscopy, chemometrics and molecular modelling	Under review	Submitted in Journal of Current Research in Food Science
5	Structural Changes of β-Casein Induced by Temperature and pH Analysed by Nuclear Magnetic Resonance, Fourier-Transform Infrared Spectroscopy, and Chemometrics	Published	Markoska, T., Daniloski, D., Vasiljevic, T. and Huppertz, T., 2021. Structural changes of β-casein induced by temperature and ph analysed by nuclear magnetic resonance, fourier-transform infrared spectroscopy, and chemometrics. Molecules, 26(24), p.7650.
6	Changes in the secondary structure of β-CN from A1 and A2 milk in regards to the pH and genetic alternative	Manuscript ready for submission	

Declaration by [candidate name]: Tatijana Markoska

Signature: Tatijana Markoska Digitally signed by Tatijana Markoska Date: 2023.03.05 18:33:31 +11'00'

Date: 05/03/2023

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 5

CHAPTER 6

CHAPTER 1: Introduction

1.1 Background

Bovine milk is biological fluid secreted in the mammary gland with important nutritive function for the neonate. The milk compromises diverse collection of constituents including proteins, fats, carbohydrates and other micronutrients. The amount of the constituents present in the milk is dependent on the stage of lactation, cow's health and age (McSweeney and Fox, 2013). The milk proteins present in milk are categorised into two main groups or caseins present in 80 % and whey proteins present in 20 % (Huppertz, 2012). Both caseins and whey proteins have vast impact on the functional, physicochemical and nutritional properties of the milk. Caseins as major protein group in the milk are classified into four individual forms including αs1-, αs2-, β- and κ-casein, which are present in milk in ratio 4.0:1.0:3.5:1.5. The caseins in milk build complex structure known as casein micelle (Horne, 2020). Each casein have its individual properties that contribute in building the micellar structure. Moreover, the structural orientation of individual caseins lead to suitable orientation of the binding sites and building interactions between caseins including hydrophobic interactions, hydrogen bonds, van der Waals interactions or for some caseins covalent interactions (Huppertz, 2012; Huppertz, Fox and Kelly, 2018). The structural orientation of caseins can be important factor to determine milk proteins' behaviour during processing. Knowing the secondary structure of the caseins is important in understanding the diverse properties that different polymorphic variant of milk adapt during milk processing.

The most common genetic variants known are A1 and A2 milk. The difference between these genotypes is in one amino acid in the polypeptide sequence of β-casein. Thus, β-casein from A1 milk have histidine in position 67 and β-casein from A2 milk have proline in the same position of the polypeptide chain (Asledottir *et al*., 2018; Daniloski *et al*., 2022a; Nguyen *et al*., 2020). Numerous studies on these two proteoforms have shown to have great effect on the dairy processing and manufacturing which affects the milk yield, gross composition and retail market (Day *et al*., 2015; Olenski *et al*., 2010; Daniloski *et al*., 2022b). The genetic polymorphism of β-casein in A1 and A2 milk was shown to have vital impact on acid or rennet induced gelation during yoghurt or cheese manufacturing (Day *et al*., 2015; Nguyen *et al*., 2018). Thus, A2 milk was shown to have low coagulation properties compared to A1 milk. The variations in the coagulation properties of these two genotypes is attributed to the presence of different amino

acid in the polypeptide chain for β-casein molecule. Thus, having only one amino acid that differ in the sequence can lead to different adaptation in the structural orientation of the molecule that can affect the milk processing.

The genotype also affects the protein digestibility in the human body by lower digestibility for A1 β-casein compared to A2 β-casein (Sebastiani *et al*., 2020; Daniloski *et al*., 2022b). The disturbed digestion results with enzymatic hydrolysis of the β-casein into smaller fragment or peptides that have opioid activity on the human body (Daniloski *et al*., 2021). These groups of peptides are known as β-casomorphins (BCMs). The fragmentation of BCMs from the β-casein polypeptide can result with multiple peptides with different length. Moreover, the primary activity of these peptides is observed in the N terminal of the sequence (Tyr-Pro-Phe) that was identified to have binding affinity towards μ-receptors in the digestive, immune and nervous system (Daniloski *et al*., 2021). The release of these peptides from β-casein molecule was observed to be more dominant for A1 β-casein type, however, some release was also observed for A2 β-casein genotype (Asledottir *et al*., 2019; Cieślińska *et al*., 2007; De Poi *et al*., 2020). Owing to the negative effect of these peptides on human health, it is essential to understand how the β-casein and some peptides behave in conditions that are crucial for digestion and milk processing. Moreover, different release level for peptides from A1 and A2 β-casein can be related to the amino acid composition that distinguish these molecules and thus the structure that the peptides adapt at certain condition. The secondary structure of β-casein and its fragment in regards to the different genotypes and conditions similar to the milk processing or digestibility have not been reported in literature so far.

The observation of the structure of β-casein and its fragments was previously observed by using different techniques including Fourier-Transform infrared spectroscopy (FTIR) and Nuclear Magnetic Resonance (NMR) (Cross *et al*., 2001; Dupont *et al*., 2013; Farrell Jr *et al*., 2001; Tsuda *et al*., 1991; Wahlgren *et al*., 1993). The FTIR provides information of the secondary structure of the molecules depicting the vibration of C=O stretching of peptides (Carbonaro and Nucara, 2010). NMR gives information of the position of individual atoms, the bonding interactions and spatial orientation of the molecule (Wüthrich, 1986). The existing studies in literature focus in observing the secondary structure components on β-casein molecule and N terminal of its polypeptide chain. However, no study exist on the conformational

transition of β-casein under conditions that are critical for milk processing and/or to determine the structure of the β-casein molecule or its fragments that are important for different processing behaviour of A1 and A2 genotype.

1.2 Research aim and objectives

The primary aim of this study is to identify the main structural differences of β-casein under various conditions and from different genotype with purpose to recognise the different behaviour of milk during processing or digestion.

The specific objectives are:

- To determine the structural behaviour of β-casomorphin 7 (BCM7) when introduced to conditions important for milk processing and/or digestion. BCM7 is the principal casomorphin released from β-casein polypeptide and main factor for possible release of smaller opioid peptides that have negative effect on human health.
- To investigate how the structure of β-casomorphin 11 from A1 and A2 milk will change when treated under similar processing conditions. BCM11 is the distinctive peptide that have one amino acid that signature A1 or A2 β-casein.
- To determine the structural transition of β-casein when treated in conditions similar to milk processing and/or digestion.
- To examine the structural characteristic of β-casein isolated from A1 and A2 milk and track the structural transition in regards to pH with aim to confirm the different coagulation properties of these genotypes in A1 and A2 milk.

1.3 Thesis outline

The thesis is organised in seven chapters. Chapter 1 present the introduction of the thesis including study background, aims and objectives and outline of the thesis. In chapter 2 presents the critical review of the literature that relates to the study with focus on the core research findings and fundamental concepts. Chapter 3 provides information of the structural transition of BCM7 under defined conditions. In chapter 4, the structure of BCM11 peptide is observed to

determine the effect of genotype and conditions on the structural components. Chapter 5 concerns the structural behaviour of β-casein in defined conditions. Chapter 6 explains structural transition of β-casein molecule form A1 and A2 genotype confirming the variations in the coagulation properties. The final chapter 7 provides conclusions of the entire work done in this project and scope for the future research.

References:

Asledottir, T., Le, T.T., Poulsen, N.A., Devold, T.G., Larsen, L.B. and Vegarud, G.E., 2018. Release of β-casomorphin-7 from bovine milk of different β-casein variants after ex vivo gastrointestinal digestion. *International dairy journal*, *81*, pp.8-11.

Asledottir, T., Picariello, G., Mamone, G., Ferranti, P., Røseth, A., Devold, T.G. and Vegarud, G.E., 2019. Degradation of β-casomorphin-7 through *in vitro* gastrointestinal and jejunal brush border membrane digestion. *Journal of dairy science*, *102*(10), pp.8622-8629.

Carbonaro, M. and Nucara, A., 2010. Secondary structure of food proteins by Fourier Transform spectroscopy in the mid-infrared region. *Amino acids*, *38*, pp.679-690.

Cieślińska, A., Kamiński, S., Kostyra, E. and Sienkiewicz-Szłapka, E., 2007. Beta-casomorphin 7 in raw and hydrolyzed milk derived from cows of alternative β-casein genotypes. *Milchwissenschaft*, *62*(2), pp.125-127.

Cross, K.J., Huq, N.L., Bicknell, W. and Reynolds, E.C., 2001. Cation-dependent structural features of β-casein-(1–25). *Biochemical Journal*, *356*(1), pp.277-286.

Daniloski, D., McCarthy, N.A., Markoska, T., Auldist, M.J. and Vasiljevic, T., 2022a. Conformational and physicochemical characteristics of bovine skim milk obtained from cows with different genetic variants of β-casein. *Food Hydrocolloids*, *124*, p.107186.

Daniloski, D., McCarthy, N.A., Huppertz, T. and Vasiljevic, T., 2022b. What is the impact of amino acid mutations in the primary structure of caseins on the composition and functionality of milk and dairy products? *Current Research in Food Science*.

Daniloski, D., McCarthy, N.A. and Vasiljevic, T., 2021. Bovine β-Casomorphins: Friends or Foes? A comprehensive assessment of evidence from *in vitro* and *ex vivo* studies. *Trends in Food Science & Technology*, *116*, pp.681-700.

Day, L., Williams, R.P.W., Otter, D. and Augustin, M.A., 2015. Casein polymorphism heterogeneity influences casein micelle size in milk of individual cows. *Journal of dairy science*, *98*(6), pp.3633-3644.

De Poi, R., De Dominicis, E., Gritti, E., Fiorese, F., Saner, S. and Polverino de Laureto, P., 2020. Development of an LC-MS method for the identification of β-casein genetic variants in bovine milk. *Food Analytical Methods*, *13*(12), pp.2177-2187.

Dupont, D., Croguennec, T., Brodkorb, A. and Kouaouci, R., 2013. Quantitation of proteins in milk and milk products. *Advanced Dairy Chemistry: Volume 1A: Proteins: Basic Aspects, 4th Edition*, pp.87-134.

Farrell Jr, H.M., Wickham, E.D., Unruh, J.J., Qi, P.X. and Hoagland, P.D., 2001. Secondary structural studies of bovine caseins: temperature dependence of β-casein structure as analyzed by circular dichroism and FTIR spectroscopy and correlation with micellization. *Food Hydrocolloids*, *15*(4-6), pp.341-354.

Huppertz, T., 2012. Chemistry of the caseins. In *Advanced Dairy Chemistry: Volume 1A: Proteins: Basic Aspects, 4th Edition* (pp. 135-160). Boston, MA: Springer US.

Huppertz, T., Fox, P.F. and Kelly, A.L., 2018. The caseins: Structure, stability, and functionality. In *Proteins in food processing* (pp. 49-92). Woodhead Publishing.

Horne, D.S., 2020. Casein micelle structure and stability. In *Milk proteins* (pp. 213-250). Academic Press.

McSweeney, P.L. and Fox, P.F. eds., 2013. *Advanced dairy chemistry: volume 1A: proteins: basic aspects*. Springer Science & Business Media.

Nguyen, D.D., Solah, V.A., Busetti, F., Smolenski, G. and Cooney, T., 2020. Application of ultra-high performance liquid chromatography coupled to high-resolution mass spectrometry (Orbitrap™) for the determination of beta-casein phenotypes in cow milk. *Food chemistry*, *307*, p.125532.

Nguyen, H.T., Schwendel, H., Harland, D. and Day, L., 2018. Differences in the yoghurt gel microstructure and physicochemical properties of bovine milk containing A1A1 and A2A2 β-casein phenotypes. *Food Research International*, *112*, pp.217-224.

Olenski, K., Kamiński, S., Szyda, J. and Cieslinska, A., 2010. Polymorphism of the beta-casein gene and its associations with breeding value for production traits of Holstein–Friesian bulls. *Livestock Science*, *131*(1), pp.137-140.

Tsuda, S., Niki, R., Kuwata, T., Tanaka, I. and Hikichi, K., 1991. 1H NMR study of casein phosphopeptide (1–25): Assignment and conformation. *Magnetic resonance in chemistry*, *29*(11), pp.1097-1102.

Sebastiani, C., Arcangeli, C., Ciullo, M., Torricelli, M., Cinti, G., Fisichella, S. and Biagetti, M., 2020. Frequencies evaluation of β-casein gene polymorphisms in dairy cows reared in Central Italy. *Animals*, *10*(2), p.252.

Wahlgren, N.M., Léonil, J., Dejmek, P. and Drakenberg, T., 1993. Two-dimensional Nuclear Magnetic Resonance study of the β-casein peptide 1–25: Resonance assignments and secondary structure. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, *1202*(1), pp.121-128.

Wüthrich, K., 1986. NMR with proteins and nucleic acids. *Europhysics News*, *17*(1), pp.11-13.

CHAPTER 2: Literature review

2A. Milk and milk proteins – general perspective

2A.1 Bovine milk

Milk is complex biological fluid secreted by the female of all mammals with principal aim to complete the nutritional requirements for the neonate. Furthermore, to meeting the neonate's nutritional requirements the milk and milk products from domesticated animals presents essential part of the human diet. As a part of the human diet milk comply with several dietary characteristics including being highly nutritional single food with pleasant and attractive flavour and it is free from toxins. The most common milk from domestic animals that is used in human diet is bovine or cow milk. The major constituent in bovine milk is water, present at ~87 %. The dry matter consists of proteins (~3.5%), lactose (~ 4.1%), fat (~ 4.5%), minerals (~0.8 %) and vitamins (~0.1 %) (O'Mahony and Fox, 2013). Milk proteins and peptides generally serve the physiological functions of milk.

2A.2 Milk proteins

Bovine milk contains approximately 3.5 % milk proteins which concentration significantly changes during the lactation period. The composition of proteins in bovine milk presents a complex mixture that is dependent on physicochemical, nutritional and functional properties of the milk. Nowadays, the milk proteins are the best characterised more than any other food proteins. The milk proteins are divided into two main groups: whey proteins that are present in milk at 20 % and caseins present at 80 % (O'Mahony and Fox, 2013).

Whey proteins are the minor group of milk proteins with many food applications due to their functionality and nutritional value. Whey proteins are comprised of mainly β-lactoglobulin (β-LG), α-lactalbumin (α-LA), immunoglobulins, and bovine serum albumin (BSA) and some minor proteins including lactoferrin, glycoproteins, transferrin, and lactollin (Kilara and Panyam, 2003). Due to important functional properties in dairy production, both β-LG and α-LA have been extensively studied for structural and functional characteristics (Brew, 2003; Sawyer, 2013).

The major milk protein group or caseins contain four different forms known as αs_1-, αs_2, β- and κ-casein (CN). In bovine milk these individual caseins are present in ratio of

4.0:1.0:3.5:1.5 for αs_1-, αs_2-, β- and κ-CN, respectively, forming one individual complex structure known as casein micelle (Huppertz, 2013). Some of the major characteristics of all four caseins are presented in Table 1.

Table 1. Main characteristics of four caseins in bovine milk (adapted from Huppertz (2013))

	Total residues	**Aromatic residues**	**pI**	**Molecular mass (Da)**
αs_1-CN	199	20	4.42	26 599
αs_2-CN	207	20	4.95	25 206
β-CN	209	14	4.65	23 973
κ-CN	169	14	5.60	19 052

The casein micelles are primary source of amino acids, calcium and phosphate for the growth requirements of mammalian neonates. The model of the casein micelle was extensively studied in the past few decades. Several models were proposed including sub-micelles, dual-binding, nanoclusters, water channels, primary casein particles and network structure (Slattery and Evard, 1973; Horne, 1998; Horne, 2020; Holt, 1992; Holt, 2004; Dalgleish and Corredig, 2012; Huppertz *et al*., 2017). The structure of casein micelle presents combination of individual caseins and colloidal calcium phosphate (Huppertz *et al*., 2018, De Kruif *et al*., 2012). The presence of calcium and phosphate is a crucial factor for micellar stabilization. The individual caseins interact and combine into nanoclusters by hydrophobic, hydrogen, electrostatic and van der Waals interactions. The nanoclusters connect with the colloidal calcium phosphate by ionic interactions stabilising the micellar structure (Lucey and Horne, 2018; De Kruif *et al*., 2012). The caseins and calcium phosphate form micellar particles with a radius of 12 nm (Schmidt, 1982, Huppertz *et al*., 2017). This complex combines 9-11 caseins including four αs_1-CN and four β-CN and either one αs_2-CN or one κ-CN (Kumosinski *et al*., 1994; Farrell *et al*., 2003). Moreover, the caseins are known to be highly phosphorylated with four phosphorylation centres for αs_1-CN, two for β-CN and two for αs_2-CN (Huppertz, 2013). The phosphorylation sites are crucial for interaction with calcium phosphate nanoclusters and stabilisation of the casein

micelles. The surface of the casein micelles is stabilised by κ-CN (Huppertz, 2013). Thus, it would be expected that the casein particles that are located at the surface of the micelles have predominately κ-CN rather than αs_2-CN, which would mean that the αs_2-CN particles are located mainly in the interior of the micelles. The casein micelles have a sponge-like structure. Thus, the micelles are highly hydrated with approximately 4 g of water per gram of protein (De Kruif and Holt, 2003). The water is unevenly distributed within the micelle by containing poorly hydrated protein rich domains and highly hydrated aqueous void spaces or water channels within the micelle (Dalgleish, 2011; Huppertz *et al*., 2017).

2A.3 Major differences between whey proteins and caseins

The whey proteins and caseins have several different structural properties. This includes different structural arrangement (secondary and tertiary structure) and amino acid composition. In milk systems, the whey proteins exist as monomers or small quaternary structures. Both α-LA and β-LG were observed to have a globular tertiary structure with define secondary structure components (Brew, 2003; Sawyer, 2013). By being homogenous, highly structured and globular proteins, whey proteins can easily be studied using different methods for structural studies including X-Ray crystallography, Nuclear Magnetic Resonance (NMR), circular dichroism (CD) and molecular dynamics (Lübke *et al*., 2002; Kuwata *et al*., 1998; Fogolari *et al*., 1998; Alexandrescu *et al*., 1993). In contrast, caseins have a highly flexible structure and complex organization in macrostructures or casein micelles. The micelles are large molecules with average diameter of 150 nm and approximately 5000 molecules. The flexible structure of caseins is due to high presence of prolines in the polypeptide chains. Proline has tendency to break the secondary structural components (Choi and Mayo, 2006). The interruption in the secondary structure components by proline is due to absence of hydrogen on the amide nitrogen and thus cannot act as donors for hydrogen bonds formation and in addition the pyrrolidine ring of the proline act as a steric constraint towards the neighbouring residues that block the formation of the structural components. Another important amino acid found in caseins is phosphoserine that is the reason for high level of phosphorylation of casein micelle. The presence of phosphoserines in caseins is crucial for metal binding (Huppertz, 2012). Thus, the most of the calcium, phosphate and zinc in the milk is found bounded to the caseins that is important for the biological

function, hydration, heat stability and structural stability of the caseins within the micelle. Despite the fact that caseins are known as molecules with flexible structure there were several secondary structure components that were identified including α-helix, β-sheets, turns and polyproline II structural components (Farrell at al., 2013). The secondary structural components that have been identified in all four caseins are presented in Table 2.

Table 2. Secondary structure components identified in four caseins (Adapted from Farrell at al., 2013)

	αs1-CN	**αs2-CN**	**β-CN**	**κ-CN**
β-Sheet	49-58	27-37	15-33	25-36
α- helix	3-5	24-32	7-25	12-17
Turns	22-31	24-31	20-30	25-28
Unspecified/random and/or PPII structures	8-23	9-22	20-25	23-24

The separation of caseins from the whey proteins and other non-caseins fractions in milk is performed using several techniques including isoelectric precipitation, centrifugation, addition of salts and calcium, membrane and gel filtration, rennet coagulation and precipitation (O'Mahony and Fox, 2013). In practice, the most used method to separate caseins is based on isoelectric precipitation by reducing the pH to approximately 4.6 that leads to coagulation of caseins. Another method is ultracentrifugation when casein micelles are separated using high-speed centrifuges (100 000g for 1 hour). This method is successful for separation of whey proteins from caseins, leading the fact that casein micelles have a large molecular mass and will participate during centrifugation while whey proteins will remain in the serum phase of the solution. The addition of calcium (rennet coagulation) and salts (e.g. $MgSO_4$, NaCl, $(NH_4)_2SO_4$) leads to precipitation of caseins along with whey proteins. The rennet coagulation is mainly applied in cheese production.

2A.4 Genetic variants of caseins in milk

Caseins in bovine milk are highly polymorphic with presence of many proteoforms (Goulding *et al*., 2020, Daniloski *et al*., 2022c). The genetic polymorphism hereditarily control the milk variants. The genetic polymorphism of the milk proteins is important for the technological properties of milk including heat stability, rennetability and protein proportions in milk (Jakoband and Puhan, 1992). So far 39 genetic variants of bovine caseins were reported including A, B, C, D, E, F, G, and H for αs_1-CN, A, B, C, and D for αs_2-CN, A1, A2, A3, A4, B, C, D, E, F, G, H1, H2, I, J and K for β-CN and A, B, B2, C, E, F1, F2, G1, G2, H, I and J for κ-CN (Gazi *et al*., 2022, Daniloski *et al*., 2022d). The genetic variations in caseins is a result of DNA mutations by the changes in the nucleotide in the protein amino sequence that can be amino acid replacement or in some cases removal (Summer *et al*., 2020). The specific amino acid sequence for individual genetic variant affects the isoelectric point and electric charge of proteins (Gai *et al*., 2021, Daniloski *et al*., 2022d). The genetic variants have different amino acid sequence which leads to different functional and physicochemical properties of milk proteins in milk and milk products.

2A.4.1 A1 and A2 β-CN genetic variants

The genetic polymorphism of β-CN was identified for the first time in Bos genus in the 20th century when three genetic variants were identified including A, B and C (Aschaffenburg, 1963). Latter the genetic variant A was confirmed to exist in four forms or A1, A2, A3 and A4 (Peterson and Kopfler, 1966; Chung *et al*., 1995). The β-CN genetic polymorphism is most investigated in research and industry (Daniloski *et al*., 2021a). The primal genotype in β-CN is known to be A2 from which A1 is developed as mutant through natural selection. The A1 variant is only identified in bovine milk, however in commercial milk both variants are present (Gai *et al*., 2021). Nowadays, the A1 and A2 are most frequent genetic variants in bovine milk (Lien, *et al*., 1999). Owing to the effect of β-CN, polymorphism on human health and dairy manufacturing the great commercial interest of A1 and A2 genotypes arise in many countries in the world (Daniloski *et al*., 2022d). The difference between A1 and A2 β-CN is in a single amino

acid in the position 67 of the polypeptide chan. More accurately, in A1 β-CN this position is occupied by histidine and in A2 β-CN in position 67 is present proline. Even though, it appears as a minor dissimilarity between these two genotypes owing the presence of 209 residues in the β-CN, the mentioned amino acid can make a big impact on casein structure and functional properties. The presence of proline in the proteins and peptides affects the secondary structure by higher favourability to polyproline II (PPII) secondary structures (Brown, and Zondlo, 2012). Thus, the structural dynamic of the β-CN and the self-assembly behaviour can be further promoted in A2 β-CN due to involvement of the additional proline in PPII formation (Raynes *et al*., 2015). Thus, in A2 milk variant were confirmed great proportions of PPII conformational structures resulted in the proline and glutamine rich regions (Daniloski *et al*., 2022b, 2022c; Syme *et al*., 2002). The involvement of β-CN in structural motif can be owned to its high hydrophobic nature (Huppertz *et al*., 2018). The C terminal (41-209) of β-CN has strong hydrophobic nature with balanced charges, while the other end of the chain or N terminal (1-40) is strongly polar and negatively charged (McCarthy *et al*., 2013). Hence, the hydrophobic regions of the β-CN are initiators of the PPII conformational motifs. In presence of other caseins, the involvement of β-CN into hydrophobic interactions would be more intensified predominately due the self-association behaviour of β-CN. The presence of additional proline in the A2 β-CN reduces the hydrophobicity of the casein and thus was observed to exist mainly in monomeric form (Raynes *et al*., 2015). Moreover, the presence of proline in the polypeptide chain favours formation of β-turns (McSweeney and Fox, 2013). The turns formation in the milk can be proline or non-proline based and 20-30 % of the present turns in the milk can result from β-CN (Huppertz, 2013; Dalgleish and Corredig, 2012). The β-CN have tendency to self-assemble due to the amphiphilic properties by comprising into structures with 15-60 β-CN molecules (Huppertz, 2013). The casein molecules are predominately held by hydrophobic interaction; however, other non-covalent interactions (van der Waals and electrostatic attractions) can further facilitate self-association (De Kruif *et al*., 2012).

In the dairy industry, the presence of A1 or A2 β-CN genetic variant in the milk affects the technological process of milk and milk products. The primary and most common method for sterilization of milk and milk products is heat treatment. The heat stability is essential for food industry ensuring safe products with prolonged shelf life. The effect of β-CN genetic variants on heat stability in milk and milk products have not been studied in depth. Moreover, it is known

that β-CN act as molecular chaperones at high temperatures by interacting with whey proteins, stabilise the aggregation, and prevent gelation in milk (Liyanaarachchi and Vasiljevic, 2018). The level of involvement of A1 milk and A1/A2 mixture into hydrophobic interactions with whey proteins at high temperature was found to be higher compared to A2 milk (Daniloski *et al*., 2022a). Thus, this confirms that A2 milk has lower coagulation properties compared to A1 milk. Moreover, the gelation properties of milk is important for the yoghurt and cheese making process. This includes the water holding capacity, curd formation, synaereses during rennet and acid coagulation. The milk coagulation by addition of rennet or acid was found to have an extended time for A2 milk compared to A1 milk (Bisutti *et al*., 2022). Moreover, the gel obtained from A2 milk is more porous, with low strength and thinner protein strains resulting with softer and creamier cheese (Mendes *et al*., 2019; Nguyen *et al*., 2018; Daniloski *et al*., 2022b).The cheese produced from these gels result with greater deformations induced by external forces and thus is more prone to breakage (Nguyen *et al*., 2018). This summarise that the presence of softer gels in A2 milk is resulted of less interaction created compared to A1 milk. The variations in the gel formations can be also guided by the difference in calcium distribution between micellar and serum phase for both genotypes.

Another difference in the technological properties between A1 and A2 milk is the formation of foam and emulsions. Since β-CN is the main constituent in the casein micelle it has been commonly used in food industry as an emulsifying and foaming agent (Chen *et al*., 2018). The β-CN from A2 milk has greater emulsion properties than β-CN from A1 milk that is due to higher solubility, less ordered structure and thus faster hydrophobic absorption to the interfacial layer (Darewicz and Dziuba, 2007). The capability of formation of foams and emulsions is important in dairy industry in production of whipped cream, ice cream, mousses to name a few.

2A.5 β-Casomorphins (BCMs)

Notwithstanding the differences in the A1 and A2, variants in regards to the effect on the technological properties in milk and milk products there are also diversities in the level of digestibility and the release of bioactive peptides during digestion. The released peptides in the human gut were found have a negative effect on human health by development of several health

conditions (Sebastiani *et al*., 2020; Daniloski *et al*., 2021a; Jarmołowska *et al*., 2007). Moreover, in the gastrointestinal system the enzymes hydrolyse the β-CN into smaller peptides or free amino acids. The liberated peptides known as β-casomorphins (BCMs) were found to have opioid activity since they can bind to the μ-receptors in the digestive, immune and nervous systems. The hydrolysed fragments in the digestive tract can be found in different polypeptide length. Some of the identified BCMs liberated from β-CN molecule (A1 or A2 are shown in Table 3).

The BCMs react with the μ-opioid receptor and are degraded by dipeptidyl peptidase IV (DPPIV) that appears to be the main degrading enzyme of BCM7 (Jarmołowska *et al*., 2007). In the encrypted casein form, these peptides do not have the opioid activity but only when hydrolysed and liberated in the digestive system (Cieślińska *et al*., 2007; Asledottir *et al*., 2019). The BCM7 was considered as a factor for development of many health conditions including cardiovascular diseases, diabetes, autism, cancer obesity to name a few (Jarmołowska *et al*., 2007; Cieślińska *et al*., 2007; Haq *et al*., 2014; Fiedorowicz *et al*., 2014); however a recent systematic review of the health implications did not support these suggestions (Daniloski *et al*., 2022d)

Table 3. Identified BCMs from A1 and A2 milk

Peptide	**Structure**	**Location in β-CN**
BCM-3	Tyr-Pro-Phe	60-62
BCM-4	Tyr-Pro-Phe-Pro	60-63
BCM-5	Tyr-Pro-Phe-Pro-Gly	60-64
BCM-6	Tyr-Pro-Phe-Pro-Gly-Pro	60-65
BCM-7	Tyr-Pro-Phe-Pro-Gly-Pro-Ile	60-66
BCM-8 (A2 β-CN)	Tyr-Pro-Phe-Pro-Gly-Pro-Ile-Pro	60-67
BCM-8 (A1 β-CN)	Tyr-Pro-Phe-Pro-Gly-Pro-Ile-His	60-67
BCM-11 (A2 β-CN)	Tyr-Pro-Phe-Pro-Gly-Pro-Ile-Pro-Asn-Ser-Leu	60-70
BCM-11 (A1 β-CN)	Tyr-Pro-Phe-Pro-Gly-Pro-Ile-His-Asn-Ser-Leu	60-70

The liberalization of BCM was found to be predominant from A1 β-CN variant compared to the A2 β-CN variant (Cieślińska *et al*., 2007; Asledottir *et al*., 2019; Lambers *et al*., 2021; De Poi *et al*., 2020). The greater amount of released BCM from A1 milk can be a result of different structural arrangement of the β-CN molecule from A1 and A2 milk. Thus, the enzymes need to access the bond to cleave the casein into smaller fragments. The presence of histidine or proline in position 67 of the A1 and A2 β-CN, respectively can be important factor for bond accessibility for yielding BCM7.

2A.5.1 Impact of structure of BCMs on functional and biological properties of milk

The BCMs isolated from β-CN molecule exhibit opioid-like activity in the human body with affinity for the μ-receptors. The structural orientation of the peptides can have significant effect on accessibility of the active sites of the peptides to react with the μ-receptors. The opioid activity is confirmed in the three amino acids of the N terminal of the fragments released from β-CN (f 60-62) or Tyr-Pro-Phe. The aromatic residues from Tyr or Phe are a crucial motif as binding site with the opioid receptors. Moreover the –OH (phenolic) group located on Tyr is essential for the opioid activity of the segmented peptides (Meisel and Fitzgerald, 2000). The positioning of the aromatic sites of the BCM3 or the *cis-trans* isomerism of the Tyr-Pro bond is essential for the binding with the μ-receptors and its activation (Mierke *et al*., 1990). Thus, the presence of Pro in position 2 of the tri-peptide is necessary for the suitable orientation of the Tyr and Pro and thus its biological activity. The Pro in the peptide chain is important factor for *cis-trans* isomerism of the X-Pro bond that affects the positioning of the active sites and accessibility of the peptide bonds for further cleavage. The Pro containing peptides reveal an equilibrium between the conformational isomers (Mierke *et al*., 1990).

In tetrapeptide - BCM4 (Tyr-Pro-Phe-Pro) *cis-trans* isomerism was also confirmed in Tyr-Pro and Phe-Pro bond (Goodman and Mierke, 1989). Thus, the *trans* isomers were dominant in both X-Pro bonds and *cis* isomers were mainly present in Tyr-Pro bond. Moreover, the different isomeric appearance of the X-Pro bond results in a different affinity for the opioid receptors. The Tyr-Pro bond have slow exchange between the two isomers confirming that the peptide does not possess clear conformational preferences (Goodman and Mierke, 1989).

The BCM5 (Tyr-Pro-Phe-Pro-Gly) peptide contains four interchanging conformational isomers (Delaet *et al*., 1991). The *cis-trans* transition of the X-Pro bond is essential mechanism for protein folding and lead to different selectivity for enzyme hydrolysis and interaction with the opioid receptors. Thus, the *trans* isomers are hydrolysed by enzymes more selectively than *cis* isomers (Fischer *et al*., 1984). Moreover, in BCM5 more favourable are *trans* isomers for both X-Pro bonds (Delaet *et al*., 1991). The conversion from one *trans* X-Pro to two *trans* X-Pro bonds in BCM is more favourable (30-55 %) than conversion to all *cis* X-pro bonds. The *trans* isomers of the Tyr-Pro bond show stacking interactions of the side chains or the aromatic residues of Tyr and Pro which is essential for the affinity of the μ-receptors (Wilkes and Schiller, 1990; Delaet *et al*., 1991).

The BCM7 (Tyr-Pro-Phe-Pro-Gly-Pro-Ile) peptide was observed to be most prominent opioid peptide liberated from either A1 or A2 β-CN (De Noni, Stuknytė and Cattaneo, 2015). The information of the structure of BCM7 in literature is limited. The presence of eight *cis-trans* isomers in the Tyr-Pro, Phe-Pro and Gly-Pro bond of the peptide chain has been confirmed (Basosi *et al*., 2001). Thus, the uneven liberalization of the BCM7 from A1 and A2 βCN can be related to the different conformational orientation of the C terminal of the peptide which is main location for enzymatic cleavage. However, the structure of BCN7 have not been revealed in details yet neither how the peptide behaves under optimal conditions for enzymatic cleavage.

The most important peptide that reveals the structural, biological and technological differences of A1 and A2 β-CN is BCM11 (Tyr-Pro-Phe-Pro-Gly-Pro-Ile-His/Pro-Asn-Ser-Leu). Moreover, the BCM11 peptide have the primary amino acids that distinguish these two genotypes. No study exist in literature revealing the conformational differences of BCM11 peptides from A1 or A2 β-CN.

2A.6 Conclusion

β-CN is principal casein responsible for the technological and digestive differences existing between A1 and A2 milk. The secondary structure of β-CN and liberated β-casomorphins can be key factor to uncover the long lasting dilemma about preferences for A1 or A2 milk. Moreover, implementing highly sophisticated methods like Nuclear Magnetic

Resonance (NMR) or Fourier-Transform infrared spectroscopy (FTIR) in structural studies of β-CN and/or its fragments from A1 and A2 milk can provide detailed information of the structural differences that exists between these two genotypes. The findings can be of great importance for industry and public by providing basic knowledge of the major structural differences of these to genotypes.

References:

Alexandrescu, A.T., Evans, P.A., Pitkeathly, M., Baum, J. and Dobson, C.M., 1993. Structure and dynamics of the acid-denatured molten globule state of. alpha.-lactalbumin: A two-dimensional NMR study. *Biochemistry*, *32*(7), pp.1707-1718.

Aschaffenburg, R., 1963. Inherited casein variants in cow's milk: II. Breed differences in the occurrence of β-casein variants. *Journal of Dairy Research*, *30*(2), pp.251-258.

Asledottir, T., Picariello, G., Mamone, G., Ferranti, P., Røseth, A., Devold, T.G. and Vegarud, G.E., 2019. Degradation of β-casomorphin-7 through *in vitro* gastrointestinal and jejunal brush border membrane digestion. *Journal of dairy science*, *102*(10), pp.8622-8629.

Basosi, R., D'Amelio, N., Gaggelli, E., Pogni, R. and Valensin, G., 2001. *cis–trans* Isomerization of β-casomorphin peptides bound to copper (II): integration of EPR and NMR studies. *Journal of the Chemical Society, Perkin Transactions 2*, (3), pp.252-257.

Bisutti, V., Pegolo, S., Giannuzzi, D., Mota, L.F.M., Vanzin, A., Toscano, A., Trevisi, E., Marsan, P.A., Brasca, M. and Cecchinato, A., 2022. The β-casein (CSN2) A2 allelic variant alters milk protein profile and slightly worsens coagulation properties in Holstein cows. *Journal of Dairy Science*, *105*(5), pp.3794-3809.

Brew, K., 2003. α-Lactalbumin. In *Advanced dairy chemistry—1 Proteins* (pp. 387-419). Springer, Boston, MA.

Brown, A.M. and Zondlo, N.J., 2012. A propensity scale for type II polyproline helices (PPII): aromatic amino acids in proline-rich sequences strongly disfavor PPII due to proline–aromatic interactions. *Biochemistry*, *51*(25), pp.5041-5051.

Chen, M., Sala, G., Van Valenberg, H.J.F., Van Hooijdonk, A.C.M., Van Der Linden, E. and Meinders, M.B.J., 2018. Foam and thin films of hydrophilic silica particles modified by β-casein. *Journal of colloid and interface science*, *513*, pp.357-366.

Choi, E.J. and Mayo, S.L., 2006. Generation and analysis of proline mutants in protein G. *Protein Engineering, Design and Selection*, *19*(6), pp.285-289.

Chung, E.R., Han, S.K. and Rhim, T.J., 1995. Milk protein polymorphisms as genetic marker in Korean native cattle. *Asian-Australasian Journal of Animal Sciences*, *8*(2), pp.187-194.

Cieślińska, A., Kamiński, S., Kostyra, E. and Sienkiewicz-Szłapka, E., 2007. Beta-casomorphin 7 in raw and hydrolyzed milk derived from cows of alternative β-casein genotypes. *Milchwissenschaft*, *62*(2), pp.125-127.

Dalgleish, D.G., 2011. On the structural models of bovine casein micelles—review and possible improvements. *Soft matter*, *7*(6), pp.2265-2272.

Dalgleish, D.G. and Corredig, M., 2012. The structure of the casein micelle of milk and its changes during processing. *Annual review of food science and technology*, *3*, pp.449-467.

Daniloski, D., McCarthy, N.A. and Vasiljevic, T., 2021a. Bovine β-Casomorphins: Friends or Foes? A comprehensive assessment of evidence from *in vitro* and *ex vivo* studies. *Trends in Food Science & Technology*, *116*, pp.681-700.

Daniloski, D., McCarthy, N.A. and Vasiljevic, T., 2022a. Impact of heating on the properties of A1/A1, A1/A2, and A2/A2 β-casein milk phenotypes. *Food Hydrocolloids*, *128*, p.107604.

Daniloski, D., McCarthy, N.A., Gazi, I. and Vasiljevic, T., 2022b. Rheological and structural properties of acid-induced milk gels as a function of β-casein phenotype. *Food Hydrocolloids*, p.107846.

Daniloski, D., McCarthy, N.A., Markoska, T., Auldist, M.J. and Vasiljevic, T., 2022c. Conformational and physicochemical characteristics of bovine skim milk obtained from cows with different genetic variants of β-casein. *Food Hydrocolloids*, *124*, p.107186.

Daniloski, D., Cunha, N.M., McCarthy, N.A., O'Callaghan, T.F., McParland, S. and Vasiljevic, T., 2021b. Health-related outcomes of genetic polymorphism of bovine β-casein variants: A systematic review of randomised controlled trials. *Trends in Food Science & Technology*, *111*, pp.233-248.

Daniloski, D., McCarthy, N.A., Huppertz, T. and Vasiljevic, T., 2022d. What is the impact of amino acid mutations in the primary structure of caseins on the composition and functionality of milk and dairy products?. *Current Research in Food Science*.

Daniloski, D., McCarthy, N.A., O'Callaghan, T.F. and Vasiljevic, T., 2022e. Authentication of β-casein milk phenotypes using FTIR spectroscopy. *International Dairy Journal*, *129*, p.105350.

Darewicz, M. and Dziuba, J., 2007. Formation and stabilization of emulsion with A1, A2 and B β-casein genetic variants. *European Food Research and Technology*, *226*(1), pp.147-152.

De Noni, I., Stuknytė, M. and Cattaneo, S., 2015. Identification of β-casomorphins 3 to 7 in cheeses and in their *in vitro* gastrointestinal digestates. *LWT-Food Science and Technology*, *63*(1), pp.550-555.

De Poi, R., De Dominicis, E., Gritti, E., Fiorese, F., Saner, S. and Polverino de Laureto, P., 2020. Development of an LC-MS method for the identification of β-casein genetic variants in bovine milk. *Food Analytical Methods*, *13*(12), pp.2177-2187.

Delaet, N.G., Verheyden, P.M., Tourwé, D. and Van Binst, G., 1991. Assignment of the 1H-nmr resonances of the four rotamers of β-casomorphin-5 in DMSO. *Biopolymers: Original Research on Biomolecules*, *31*(12), pp.1409-1416.

De Kruif, C.G. and Holt, C., 2003. Casein micelle structure, functions and interactions. In *Advanced dairy chemistry—1 proteins* (pp. 233-276). Springer, Boston, MA.

De Kruif, C.G., Huppertz, T., Urban, V.S. and Petukhov, A.V., 2012. Casein micelles and their internal structure. *Advances in colloid and interface science*, *171*, pp.36-52.

Farrell, H.M., Brown, E.M. and Malin, E.L., 2013. Higher order structures of the caseins: A paradox?. In *Advanced dairy chemistry* (pp. 161-184). Springer, Boston, MA.

Farrell, H.M., Brown, E.M., Hoagland, P.D. and Malin, E.L., 2003. Higher order structures of the caseins: a paradox?. In *Advanced Dairy Chemistry—1 Proteins* (pp. 203-231). Springer, Boston, MA.

Fiedorowicz, E., Kaczmarski, M., Cieślińska, A., Sienkiewicz-Szłapka, E., Jarmołowska, B., Chwała, B. and Kostyra, E., 2014. β-casomorphin-7 alters μ-opioid receptor and dipeptidyl peptidase IV genes expression in children with atopic dermatitis. *Peptides*, *62*, pp.144-149.

Fischer, G., Bang, H. and Mech, C., 1984. Determination of enzymatic catalysis for the *cis-trans*-isomerization of peptide binding in proline-containing peptides. *Biomedica biochimica acta*, *43*(10), pp.1101-1111.

Fogolari, F., Ragona, L., Zetta, L., Romagnoli, S., De Kruif, K.G. and Molinari, H., 1998. Monomeric bovine β-lactoglobulin adopts a β-barrel fold at pH 2. *FEBS letters*, *436*(2), pp.149-154.

Gai, N., Uniacke-Lowe, T., O'Regan, J., Faulkner, H. and Kelly, A.L., 2021. Effect of protein genotypes on physicochemical properties and protein functionality of bovine milk: a review. *Foods*, *10*(10), p.2409.

Gazi, I., Johansen, L.B. and Huppertz, T., 2022. Heterogeneity, Fractionation, and Isolation.

Jakob, E. and Puhan, Z., 1992. Technological properties of milk as influenced by genetic polymorphism of milk proteins—A review. *International dairy journal*, *2*(3), pp.157-178.

Goodman, M. and Mierke, D.F., 1989. Configurations of morphiceptins by proton and carbon-13 NMR spectroscopy. *Journal of the American Chemical Society*, *111*(10), pp.3489-3496.

Goulding, D.A., Fox, P.F. and O'Mahony, J.A., 2020. Milk proteins: An overview. *Milk proteins*, pp.21-98.

Haq, M.R.U., Kapila, R., Sharma, R., Saliganti, V. and Kapila, S., 2014. Comparative evaluation of cow β-casein variants (A1/A2) consumption on Th2-mediated inflammatory response in mouse gut. *European journal of nutrition*, *53*(4), pp.1039-1049.

Holt, C., 2004. An equilibrium thermodynamic model of the sequestration of calcium phosphate by casein micelles and its application to the calculation of the partition of salts in milk. *European Biophysics Journal*, *33*(5), pp.421-434.

Holt, C., 1992. Structure and stability of bovine casein micelles. *Advances in protein chemistry*, *43*, pp.63-151.

Horne, D.S., 2020. Casein micelle structure and stability. In *Milk proteins* (pp. 213-250). Academic Press.

Horne, D.S., 1998. Casein interactions: casting light on the black boxes, the structure in dairy products. *International Dairy Journal*, *8*(3), pp.171-177.

Huppertz, T., Fox, P.F. and Kelly, A.L., 2018. The caseins: Structure, stability, and functionality. In *Proteins in food processing* (pp. 49-92). Woodhead Publishing.

Huppertz, T., 2012. Chemistry of the caseins. In *Advanced dairy chemistry* (pp. 135-160). Springer, Boston, MA.

Huppertz, T., Gazi, I., Luyten, H., Nieuwenhuijse, H., Alting, A. and Schokker, E., 2017. Hydration of casein micelles and caseinates: Implications for casein micelle structure. *International Dairy Journal*, *74*, pp.1-11.

Jann, O., Ceriotti, G., Caroli, A. and Erhardt, G., 2002. A new variant in exon VII of bovine β-casein gene (CSN2) and its distribution among European cattle breeds. *Journal of animal breeding and genetics*, *119*(1), pp.65-68.

Jarmołowska, B., Bielikowicz, K., Iwan, M., Sidor, K., Kostyra, E. and Kaczmarski, M., 2007. Serum activity of dipeptidyl peptidase IV (DPPIV; EC 3.4. 14.5) in breast-fed infants with symptoms of allergy. *Peptides*, *28*(3), pp.678-682.

Kilara, A. and Panyam, D., 2003. Peptides from milk proteins and their properties. *Critical Reviews in Food Science and Nutrition*, *43*(6), pp.607-633.

Kumosinski, T.F., King, G. and Farrell, H.M., 1994. An energy-minimized casein submicelle working model. *Journal of protein chemistry*, *13*(8), pp.681-700.

Kuwata, K., Hoshino, M., Era, S., Batt, C.A. and Goto, Y., 1998. α→ β transition of β-lactoglobulin as evidenced by heteronuclear NMR. *Journal of molecular biology*, *283*(4), pp.731-739.

Lambers, T.T., Broeren, S., Heck, J., Bragt, M. and Huppertz, T., 2021. Processing affects beta-casomorphin peptide formation during simulated gastrointestinal digestion in both A1 and A2 milk. *International Dairy Journal*, *121*, p.105099.

Lien, S., Kantanen, J., Olsaker, 1., Holm, L.E., Eythorsdottir, E., Sandberg, K., Dalsgard, B. and Adalsteinsson, S., 1999. Comparison of milk protein allele frequencies in Nordic cattle breeds. *Animal Genetics*, *30*(2), pp.85-91.

Liyanaarachchi, W.S. and Vasiljevic, T., 2018. Caseins and their interactions that modify heat aggregation of whey proteins in commercial dairy mixtures. *International Dairy Journal*, *83*, pp.43-51.

Lucey, J.A. and Horne, D.S., 2018. Perspectives on casein interactions. *International Dairy Journal*, *85*, pp.56-65.

Lübke, M., Guichard, E., Tromelin, A. and Le Quéré, J.L., 2002. Nuclear Magnetic Resonance spectroscopic study of β-lactoglobulin interactions with two flavor compounds, γ-decalactone and β-ionone. *Journal of agricultural and food chemistry*, *50*(24), pp.7094-7099.

Meisel, H. and Fitzgerald, R.J., 2000. Opioid peptides encrypted in intact milk protein sequences. *British Journal of Nutrition*, *84*(S1), pp.27-31.

Mierke, D.F., Nößner, G.E.R.H.A.R.D., Schiller, P.W. and Goodman, M., 1990. Morphiceptin analogs containing 2-aminocyclopentane carboxylic acid as a peptidomimetic for proline. *International Journal of Peptide and Protein Research*, *35*(1), pp.35-45.

Mendes, M.O., de Morais, M.F. and Rodrigues, J.F., 2019. A2A2 milk: Brazilian consumers' opinions and effect on sensory characteristics of Petit Suisse and Minas cheeses. *LWT*, *108*, pp.207-213.

McCarthy, N.A., Kelly, A.L., O'Mahony, J.A. and Fenelon, M.A., 2013. The physical characteristics and emulsification properties of partially dephosphorylated bovine β-casein. *Food chemistry*, *138*(2-3), pp.1304-1311.

McSweeney, P.L. and Fox, P.F. eds., 2013. *Advanced dairy chemistry: volume 1A: proteins: basic aspects*. Springer Science & Business Media.

Nguyen, H.T., Schwendel, H., Harland, D. and Day, L., 2018. Differences in the yoghurt gel microstructure and physicochemical properties of bovine milk containing A1A1 and A2A2 β-casein phenotypes. *Food Research International*, *112*, pp.217-224.

O'mahony, J.A. and Fox, P.F., 2013. Milk proteins: Introduction and historical aspects. In *Advanced dairy chemistry* (pp. 43-85). Springer, Boston, MA.

Peterson, R.F. and Kopfler, F.C., 1966. Detection of new types of β-casein by polyacrylamide gel electrophoresis at acid pH: A proposed nomenclature. *Biochemical and Biophysical Research Communications*, *22*(4), pp.388-392.

Raynes, J.K., Day, L., Augustin, M.A. and Carver, J.A., 2015. Structural differences between bovine A1 and A2 β-casein alter micelle self-assembly and influence molecular chaperone activity. *Journal of dairy science*, *98*(4), pp.2172-2182.

Sawyer, L., 2013. β-Lactoglobulin. In *Advanced dairy chemistry* (pp. 211-259). Springer, Boston, MA.

Schmidt, D.G., 1982. Association of caseins and casein micelle structure. *Developments in dairy chemistry*, *1*, pp.61-86.

Sebastiani, C., Arcangeli, C., Ciullo, M., Torricelli, M., Cinti, G., Fisichella, S. and Biagetti, M., 2020. Frequencies evaluation of β-casein gene polymorphisms in dairy cows reared in Central Italy. *Animals*, *10*(2), p.252.

Slattery, C.W. and Evard, R., 1973. A model for the formation and structure of casein micelles from subunits of variable composition. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, *317*(2), pp.529-538.

Summer, A., Di Frangia, F., Ajmone Marsan, P., De Noni, I. and Malacarne, M., 2020. Occurrence, biological properties and potential effects on human health of β-casomorphin 7: Current knowledge and concerns. *Critical Reviews in Food Science and Nutrition*, *60*(21), pp.3705-3723.

Syme, C.D., Blanch, E.W., Holt, C., Jakes, R., Goedert, M., Hecht, L. and Barron, L.D., 2002. A Raman optical activity study of rheomorphism in caseins, synucleins and tau: New insight into the structure and behaviour of natively unfolded proteins. *European Journal of Biochemistry*, *269*(1), pp.148-156.

Wilkes, B.C. and Schiller, P.W., 1990. Conformation-Activity relationships of cyclic dermorphin analogues. *Biopolymers: Original Research on Biomolecules*, *29*(1), pp.89-95.

Chapter 2: Literature review

2B. Unravelling conformational aspects of milk protein structure—Contributions from Nuclear Magnetic Resonance studies

---

Review article published in Foods 2020

DocuSign Envelope ID: C7AFBA28-8EE3-42F6-BADC-B4D1B5ADAD78

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book: Unravelling Conformational Aspects of Milk Protein Structure—Contributions from Nuclear Magnetic Resonance Studies.

Surname: Markoska First name: Tatijana

Institute: Institute for Sustainable Industries and Liveat Candidate's Contribution (%): 85

Status:

Accepted and in press: ✔ Date: 13/08/2020

Published: ✔ Date: 16/08/2020

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Tatijana Markoska Digitally signed by Tatijana Markoska Date: 2023.03.05 18:50:48 +11'00' 05/03/2023

Signature **Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne)
02475D (Sydney), RTO 3113

DocuSign Envelope ID: C7AFBA28-6EE3-42F6-BA0C-B4D1B5ADAD78

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

College of Health and Biomedicine, Victoria University, Werribee Campus, Melbourne, Victoria, Australia.

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Tatijana Markoska	85%	Concept development and manuscript writing and revision	[redacted]	30/3/23
Todor Vasiljevic	10%	Concept development, manuscript editing and journal submission		30/03/2023
Thom Huppertz	5%	Concept development and manuscript editing		3/30/2023

Updated: September 2019

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne),
02475D (Sydney), RTO 3113

foods

MDPI

Review

Unravelling Conformational Aspects of Milk Protein Structure—Contributions from Nuclear Magnetic Resonance Studies

Tatijana Markoska [1], Todor Vasiljevic [1] and Thom Huppertz [1,2,3,*]

[1] Advanced Food Systems Research Unit, Institute for Sustainable Industries and Liveable Cities, College of Health and Biomedicine, Victoria University, Melbourne VIC 8001, Australia; tatijana.markoska@live.vu.edu.au (T.M.); todor.vasiljevic@vu.edu.au (T.V.)

[2] FrieslandCampina, 3818 LE Amersfoort, The Netherlands

[3] Food Quality and Design Group, Wageningen University and Research, 6808 WG Wageningen, The Netherlands

* Correspondence: thom.huppertz@wur.nl

Received: 25 July 2020; Accepted: 13 August 2020; Published: 16 August 2020

Abstract: Changes in the molecular structure and association of milk proteins lead to many desirable (under controlled conditions) or undesirable characteristics of dairy products. Several methods have been used to study the structure of milk proteins and changes therein in different environments. Whey proteins are an excellent model for secondary structure studies using circular dichroism (CD), Fourier-transform infrared spectroscopy (FTIR) and tertiary structure studies using X-ray crystallography and nuclear magnetic resonance (NMR). However, caseins, the most abundant protein class in milk, are far more difficult to characterize. The tertiary structure of caseins cannot be observed by X-ray crystallography due to the inability to crystallize caseins. However, NMR is an appropriate approach for structural elucidation. Thus far, NMR was applied on specific peptides of individual caseins of the molecules including phosphoserine centers and colloidal calcium phosphate. The literature focuses on these parts of the molecule due to its importance in building the sub-unit particles involving individual caseins and calcium phosphate nanoclusters. This review focuses on present structural studies of milk proteins using NMR and their importance in dairy processing.

Keywords: NMR; milk protein; protein structure; casein; whey protein

1. Introduction

Milk is a biological fluid secreted by the mammary gland with the primary function to complete the nutritional requirements of neonates. In the dairy industry, milk is processed for maintaining safety and extending the shelf-life but is also further processed for production of different dairy products and ingredients. Milk proteins play an essential part in obtaining high quality products when appropriate processes are applied. The structures of milk proteins often undergo transformations during manufacturing processes, which may influence the quality of the final product [1]. The distribution of the amino acids in the polypeptide chain in the protein molecule is of great importance to their conformational arrangements and consequent functionality during dairy processing.

The proteins in milk are classified in two main groups, i.e., the whey proteins and the caseins and these groups are likely the most studied food proteins to date. The major whey proteins in milk are α-lactalbumin (α-LA) and β-lactoglobulin (β-LG) and considering their nutritional importance, molecular size (relatively small globular proteins) and heat sensitivity, they have been frequently studied for structural elucidation and changes therein [1–5]. The caseins occur mainly in the form of casein micelles and their complex structure has been intriguing dairy scientists for decades. The

structure of the micelles consists of four casein types, i.e., α_{s1}-, α_{s2}- and β-casein located primarily in the interior and κ-casein located on the exterior of the micelles. The primary role of κ-casein is the stabilization of the micelles against aggregation. The internal structure is stabilized by calcium phosphate clusters, which bind the phosphoserine regions of caseins and thus maintain the internal structure of the micelle [6]. During processing, these internal and external interactions may be altered, leading to structural modifications of individual caseins and casein micelle structure.

Understanding the complexity of casein arrangement in the micelle is essential for the process control. Even though the substructure of the micelle was extensively studied [7–10], some aspects of the internal and external organization, especially its changes during processing, remain unclear. Changes in the micellar equilibrium take place during processing steps including addition of acids, heat treatment, cooling, pressuring, renneting, addition of cations etc. [1]. Certain physio-chemical changes that take place during some of the dominant technological processes are illustrated in Figure 1. Modified conditions affect the micelle in a different manner and intensity. Understanding these changes is important for achieving consistent product quality and process control.

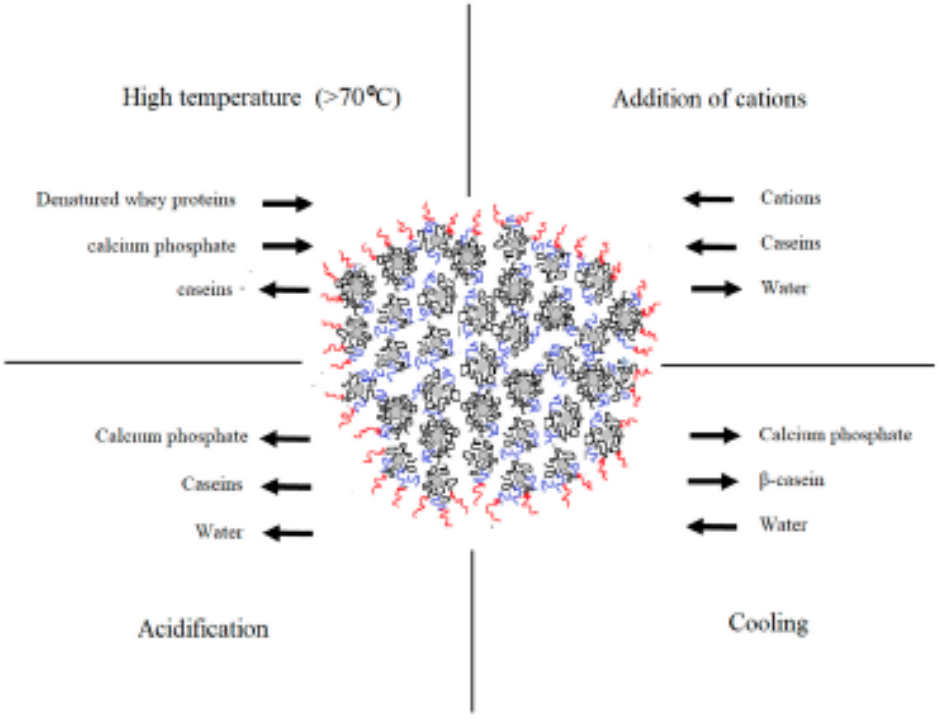


Figure 1. Changes in the native equilibrium of casein micelle in particular physio-chemical conditions during dairy processing (adapted from Gaucheron [11]).

The functional properties of milk proteins are dependent on their structural organizations. Thus, any changes in the protein structure induced by a modified environment in the milk systems affect their properties. As previously discussed, casein micelle has a complex structure where each individual casein has an important function to maintain its stability. The functional properties of casein micelle are closely related to its surface and interior properties [1]. Hence, the C-terminal of κ-casein, including the casein macropeptide (CMP), comprising of residues 106–169, is located on the exterior of the micelle. This region contains a negative charge that arises from phosphorylation and glycosylation which provide steric stabilization and thus prevent the close approach of other micelles [6]. In addition, the amino acids of κ-casein that are responsible for disulphide interactions (Cys_{11} and Cys_{88}) are located close to the surface of the micelle. However, this equilibrium can be disrupted when different processing conditions are applied which result in structural reorganization of milk proteins.

In addition, the macropeptide is not so dense and thus when the milk equilibrium is disrupted, individual proteins can enter or exit the micelle using the micellar water channels [9]. Some of these modifications are a dissociation of β-casein from the micelle during cooling [12], causing the liberalization of CMP as a result of bond cleavage (105–106) by chymosin [13] and penetration of denatured β-LG inside the κ-casein layer and formation of covalent bonds with free Cys residues during milk pre-warming [14]. The interior of the micelle also undergoes structural reorganizations that affect protein functionality, including post-coagulation changes during cheese making [1] or dissociation of caseins and insolubilization of calcium phosphate at high temperatures [14]. All these modifications are the result of protein unfolding, exposing of their active sides and involvement of different interactions which may lead to aggregation and gelation. The structural changes of milk proteins lead to important outcomes in the processing performance of the final milk product.

To elucidate milk protein structure and changes therein, the most frequently applied techniques include circular dichroism (CD) and Fourier-transform infrared spectroscopy (FTIR), which focus on various elements of the secondary structure [15]. CD records the signal arising from the peptide bond (170–240 nm) and aromatic residues (260–320 nm) of the proteins. Being a non-destructive and suitable for aqueous solutions, CD has been successfully applied on structural studies on milk proteins [16–18]. However, the method is more reliable for observation of helical structures than β-sheets, and therefore more applicable for globular proteins such as β-LG and α-LA [19]. In addition, the applicability of CD is also limited by the fact that only a limited number of buffers can be used in sample preparation. Caseins have open and unordered conformation and this can bring difficulties in the application of CD [20].

FTIR has been successfully applied in secondary structure studies of milk proteins using an ATR (attenuated total reflectance) cell depicting C=O and N-H stretching of the molecules [20–22]. However, FTIR only provides information of the presence of elements of the secondary structure in the observed system without clear differentiation of which part of the molecule is responsible for it. X-ray crystallography uses the three-dimensional density pattern of the molecule in forming an X-ray pattern and can record the atomic distance with accuracy of 0.1–0.2 Å using resolution of 1.5–2 Å range. However, this method requires the protein molecule to be crystallized. Hence, globular proteins, including α-LA and β-LG, are valid structural models for X-ray crystallography and their three-dimensional structure was successfully established at 1.8 Å resolution using this technique [23]. Caseins, due to their high surface hydrophobicity, heterogenicity and structural flexibility cannot be crystallized and thus are not suitable for structural studies by X-ray crystallography [24].

Methods such as FTIR can show the changes in the proteins as a result of processing or altered environments, but a more sophisticated approach, such as NMR, is needed to evaluate the reasons of the structural rearrangements of individual proteins and the possibility of their control. However, the complexity in theoretical and technical manipulation, including the high instrumental cost, often brings a hesitation to its selection. In comparison to X-ray crystallography, NMR can provide structural information almost under any condition applicable with a minimal sample preparation. Hence, this makes NMR an important technique for studying casein conformation.

Current knowledge describes only a few NMR structural studies on individual caseins, specific peptide fragments to be exact, concentrating only on their phosphoserine regions due to their importance in interactions with colloidal calcium phosphate and creation of a micellar skeleton [16,25–34]. Caseins possess a rather random structure that takes flexible conformations and thus are difficult models for structural studies. On the other hand, NMR has been shown to be a reliable technique to ascertain structural properties and modifications of many proteins. NMR provides information of the position of individual atoms which can present the spatial orientation of the molecule. This technique was proven to be reliable for structural studies on whey proteins, providing detailed information of the position of individual protons and, hence, the amino acid location in the polypeptide chain [35]. At present, the protein data bank (PDB) contains 10,520 resolved structures using the NMR method out of the 165,957 reported solved conformations of proteins (rcsb.org).

Many studies have been reported in the literature that include structural elucidation of milk proteins using NMR, including predominately conformational structures of whey proteins and only some studies on caseins structure. The focus of this review is to elaborate on the importance of understanding conformational changes of milk proteins during processing and consequently their impact on the quality of the products, and how these changes can be ascertained with emphasis on suitability and application of NMR as a cutting-edge approach in elucidating protein conformations.

2. NMR Approach for Structural Elucidation of Proteins

Proteins are composed of amino acids organized in a polypeptide chain, which can be in a folded or flexible organization in order to function properly. The tertiary structure of proteins has become relatively easy to predict based on a known amino acid sequence and complete spatial organization, which can be established by multidimensional NMR experiments. Consequently, high resolution NMR spectroscopy can be used to observe structural characteristics of many proteins, including milk proteins. The most studied nuclei are ^{1}H, ^{15}N and ^{31}P. Each nucleus gives a specific chemical shift in the spectral view. For a complex molecular structure such as that of a protein, one-dimensional NMR is not sufficient for identification of structural features; therefore, higher dimensions are applied. The interpretation of these higher-order spectra can be challenging and dependent on the molecular weight of the protein and the method used. NMR studies on proteins with molecular weight up to 20 kDa had become a routine due to improvements in technology. This refers to a greater availability of experiments in higher dimensions (three- and four-dimensional methods), improvement in the radiofrequency, field strength, isotopic labelling, different NMR probes and software [36].

For NMR studies of large molecules (20–50 kDa), including proteins, molecular/isotopic labelling techniques are available. The development of the NMR spectroscopy produced new techniques for optimal labelling of proteins. Many approaches for isotopic labelling have been reported, which makes the selection of an appropriate method difficult. Tugarinov et al. [37] proposed four approaches for successful labelling of proteins, i.e.,:

(1) For a backbone assignment, including $^{13}C\beta$ nuclei, the best approach is full labelling of ^{15}N, ^{2}H and ^{13}C samples obtained from D_2O based growths. The measurements are performed on protein dispersed in H_2O after $^{2}H > {}^{1}H$ exchange;

(2) For Leu, Ile δ_1 and Val methyl groups and measurements of $^{3}J_{CyCO}$ scalar coupling and nuclear Overhauser effect (NOE) connectivity (NH-CH_3; HN-HN distance), the most appropriate labelling procedure is considered to be linearized ^{13}C spin system including ((U-^{15}N,^{2}H,^{13}C), Leu, Val ($^{13}CH_3$, $^{12}CD_3$), Ileδ_1($^{13}CH_3$));

(3) A methyl labelling scheme similar to step 2, but including different carbon positions ^{12}C ((U-^{15}N,^{2}H), Leu, Val ($^{13}CH_3$, $^{12}CD_3$), Ileδ_1 ($^{13}CH_3$)), should be used for measurements of $^{3}J_{CyN}$ coupling and NOE connectivity (CH_3-CH_3);

(4) Methyl labelling as $^{13}CHD_2$- labelled proteins for detecting methyl ^{13}C relaxation rates.

In order to achieve a complete structural view of the observed protein by NMR, several steps need to be followed. These include chemical shift assignment, nuclear Overhauser effect (NOE) coupling and relaxation measurements for internal mobility of the proteins [38]. The first step of structural identification presents an individual chemical shift of the observed nuclei. For milk proteins, the most observed nuclei are ^{1}H, ^{15}N and ^{31}P, which produce characteristic chemical shifts, measured in ppm. Moreover, when the chemical shift assignment is available, appropriate combinations of NMR experiments can produce information of the structure and dynamics of the observed protein.

For sequential connectivity among the nuclei in proteins, some of the most used experiments are ^{1}H-^{15}N and ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC), ^{1}H-^{1}H correlation spectroscopy (COSY) and ^{1}H-^{1}H total correlation spectroscopy (TOCSY) experiments [39]. The ^{1}H-^{15}N HSQC and ^{1}H-^{13}C HSQC methods record only one signal for every amino acid or the backbone N-H and C-H chemical shift, respectively [40]. This excludes the Pro residues, due to their lack of amide proton in

the molecule. The COSY spectra give information of the position of α-protons in the molecule [41]. TOCSY has an important role for localization of amino acids in the spectra predominately in the amino/fingerprint region where every amino acid appears in a specific pattern [42]. These methods give information of the neighboring atom that interacts through various bonds (Figure 2a). However, this approach only provides information about which nuclei are connected through bonds and is not sufficient for structural studies. The best performance is obtained when all experiments are recorded in H_2O/D_2O solution in order to reduce the rapid exchangeability of the amide protons with the solvent [43].

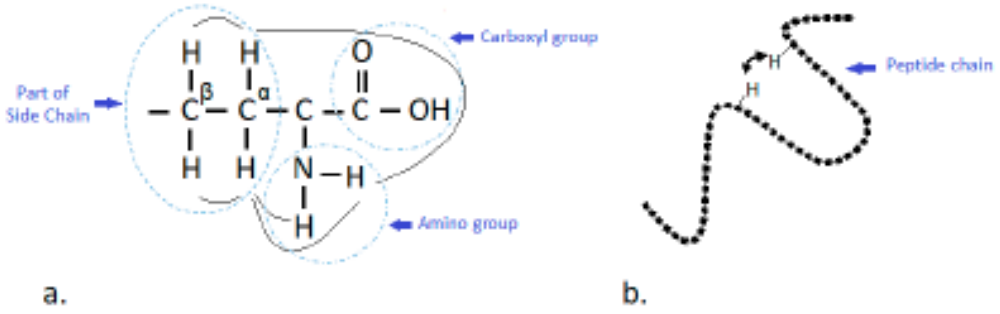


Figure 2. Through bound (scalar) coupling (**a**); through space coupling—nuclear Overhauser effect (NOE) (**b**).

The second step and most important part of the structure elucidation is a NOE experiment. The spectra from nuclear Overhauser effect spectroscopy (NOESY) are obtained through space coupling between protons which are located in the spatial distance less than 5 Å [35] (Figure 2b). NOE spectra were used for the first time for sequence specific resonance assignments for proteins by Wüthrich [35] providing two types of NOE distance among protons including a strong NOE with an upper distance limit of ≤2.5 Å and a weak NOE distance of ≤4 Å. The quantitative distance measurements of the NOE experiments depend on proton/proton distance (r) and rotation correlation time (τ_c) (Equation (1)) [38].

$$\text{NOE} \propto \frac{1}{(r^6)} \times f(\tau_c) \quad (1)$$

Wüthrich et al. [43] proposed three groups of distance constraints obtained from the NOESY spectra. The first group includes characterization of the secondary structure of the protein by closely spaced nuclei or backbone H_α and H_β atoms. The second group includes distance constraints in the sequence between H_α and H_β nuclei from further apart residues in the polypeptide chain which is important for observation of extended polypeptide structures and β-sheets or α-helix. The final group includes the distance constraints among the hydrogen atoms on the side chains which characterize the tertiary structure of the molecule. The connectivity between neighboring amino acids is established by $d_{AB}(i,j)$ presenting the distance between hydrogen atoms (A and B) located in position i and j, respectively. Depending on the spectral region observed, the NOE can provide information of distance between protons in amide region $d_{NN}(i,j)$, α-protons and amide $d_{\alpha N}(i,j)$, β-protons and amide $d_{\beta N}(i,j)$, α-α-protons $d_{\alpha\alpha}(i,j)$ and α-β-protons $d_{\alpha\beta}(i,j)$ [44]. Moreover, the distance between protons located at the nearby amino acid in the sequence is presented as a sequential distance, e.g., $d_{\alpha N}(i, i+1; i+2, i+3 \ldots)$ [44]. The distance constraints present between the backbone and the H^β of the residues, which are closely positioned in the polypeptide chain, can give information of the specific secondary structure [43]. Moreover, the distance between the backbone and the H^β residues that are further apart in the polypeptide chain can give information of the existence of super secondary structures, i.e., β-sheet or α-helix. The cross peaks in the NOESY spectra give information of the distance between two observed protons in the peptides and thus provide information of the existence of possible structures. A simple presentation of the through space connectivity in an α- helix is presented in Figure 3.

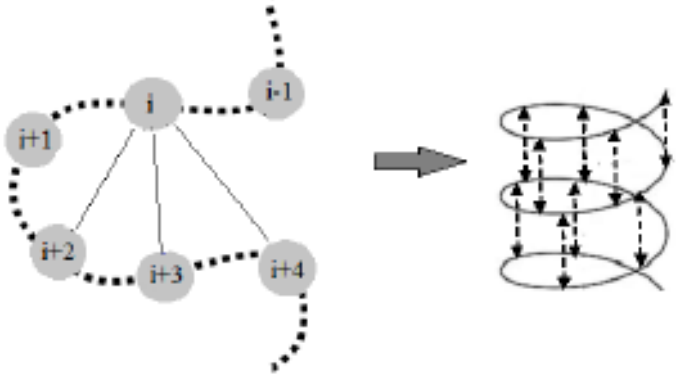


Figure 3. Regular NOE patterns for secondary stricture determination of α-helix.

NOE experiments also give information of ϕ and ψ dihedral angle conformations via spin–spin coupling constraints or J-coupling constraints by the Karplus relationship [45]. The ϕ and ψ dihedral angles arise from different combinations of steric effects that took place between the residues of the same amino acid and between the side chains from different amino acids in the polypeptide chains. These angles also result from aroused hydrogen bonds during secondary structure interactions [46]. This information is an essential parameter for backbone conformation determination, including α-helices and β-sheets.

The final step or relaxation measurements observe the recovery time for a population of spins in the observed nucleus after turning off the induced radiofrequency. In dairy products, NMR relaxation studies are mainly used to observe the water mobility and water holding capacity of the products. This includes studies that observe the state of free water and water attached to the protein matrix in the dairy products. In addition, for hydrogen as the most abundant nuclei in the organic components, two observing parameters are being obtained including spin–lattice relaxation time with a time constant T_1 and spin–spin relaxation time with a time constant T_2. Proton relaxation studies were explained in detail by Mariette [47] giving an overview of different applications and examples of NMR relaxation studies, diffusion coefficient measurements and magnetic resonance imaging (MRI) on dairy products. The relaxation measurement studies observe the water distribution and dynamics in the milk products, mainly with regard to the interactions among proteins and water in milk.

3. NMR Studies on Structure of Milk Proteins

3.1. Whey Proteins

When it comes to NMR studies on milk proteins, whey proteins have been studied more extensively than caseins. The main focus has been on understanding their thermal instability with α-LA and β-LG, as the predominant whey proteins, being the most unstable. As globular proteins, α-LA and β-LG have served as absolute structural models for NMR studies. Complete structural assignments have been reported in literature using predominately 2D, 3D NMR and isotopic labelling. Structural characterization and the NMR methods used are further discussed in this review.

3.1.1. α-Lactalbumin

α-LA exists as two domains including A domain, organized in four helixes and 3_{10}-helix, and B domain consisted of triple anti-parallel β-sheets and 3_{10}-helix [48]. α-LA was observed to exist in a molten globule state during the transition process under denaturing conditions in the early stages of refolding [49]. The NMR spectra of the molten globule of α-LA give broad linewidths and poor chemical shifts [50–52]. The broad linewidths of the NMR spectrum correspond to a compact and folded structure of the molecule [50] and well-defined sharp peaks correspond to unfolded parts of

the protein [51,53]. Predominant NMR methods for studying the molten globule structure are N-H HSQC [51,54] and NOE connectivity for 2,3 D methods [52,53,55].

The structure of α-LA was studied by Alexandrescu et al. [52,56]. They used 2D NMR methods in combination with photochemically-induced dynamic nuclear polarization (PCIDNP) to observe the aromatic rings and cluster formation in amino acids in the molecule. The sequence-specific assignments among protons with aromatic rings were detected based on NOE connectivity [56]. In their next work they applied 2D NMR spectroscopy to observe the acid-denatured molten globule structure of α-LA (101–110) by using the sequential NOE connectivity for structural assignment and observed nonspecific random coil structure of the molten globule state of α-LA [52]. The unfolding transition of the molten globule of α-LA was observed in higher dimension by Schulman et al. [51] using ^{15}N-labelled human α-LA, as well as 2D and 3D-edited NMR methods. This study confirms the unfolding transition of the molten globule of α-La in the presence of a denaturant, resulting in a non-cooperative collapsing of the polypeptide chain.

Later, Forge et al. [3] used NMR in combination with CD and mass spectroscopy to observe a refolding of α-LA after chemical denaturation. They observed that the addition of Ca^{2+} after denaturation promotes conversion of the protein from a partially folded state to the native state, which is in line with studies reported by Berliner et al. [57] showing the importance of cation-interactions for α-LA. The study was performed using fully labelled α-LA protein for real-time NMR experiments (kinetics measurements) and 3D NMR methods (HSQC-TOCSY and HSQC-NOESY) for sequential assignment. Complete sequential assignments of the chemical shift were performed, and then deposited in the BioMagRes data bank (code-4332). They identified the presence of two state processes of association for refolding from the denatured to native state. Moreover, α-LA has served as a model for NMR studies, the structural studies for sequential assignments of this whey protein that are listed in Table 1.

Table 1. NMR structural studies on α-LA and β-LG, including the performed NMR methods and additional methods used. Abbreviations: DQF-COSY—double quantum-filtered correlated spectroscopy; RELAY—relayed correlation spectroscopy NOESY—nuclear Overhauser effect spectroscopy; PCIDNP—photochemically induced dynamic nuclear polarization; TOCSY—total correlated spectroscopy; ROESY—rotating frame Overhauser Effect spectroscopy; HSQC—heteronuclear single quantum coherence; CD—circular dichroism; DYANA—dynamics algorithm for NMR applications; ARIA—ambiguous restraints for iterative assignment; PCA—principal component analysis; WATERGATE—water suppression by gradient tailored excitation.

Protein	NMR Methods Used	Additional Methods	Reference
α-LA	DQF-COSY, RELAY, NOESY	PCODNP	[56]
α-LA	DQF-COSY, NOESY, RELAY		[52]
α-LA	DQF-COSY, TOCSY, NOESY, ROESY, presaturation	CD	[58]
α-LA	3D: TOCSY-HSQC, NOESY-HSQC, TOCSY, NOESY	CD	[3]
α-LA	3D: ^{15}N-edited NOESY-HSQC, COSY		[53]
β-LG	DQF-COSY, TOCSY, NOESY	CD	[17]
β-LG	DQF-COSY, TOCSY, NOESY		[18]
β-LG	^{13}C, ^{15}N-protein labelling 2D: ^{15}N-HSQC, TOCSY 3D: CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, HCCH-TOCSY, CCH-TOCSY, $(H^{\beta})C^{\beta}(C\gamma C^{\delta})H^{\delta}$ $H^{\beta}(C^{\beta}C\gamma C^{\delta})H^{\delta}$, ^{1}H-^{1}H-^{15}N NOESY	CD, X-ray scattering	[59]
β-LG	DQF-COSY, TOCSY, NOESY	Structure calculation (DYANA)	[60]

Table 1. *Cont.*

Protein	NMR Methods Used	Additional Methods	Reference
β-LG	^{15}N,^{13}C-labelled proteins 2D: ^{1}H-^{15}N HSQC 3D: CBCA(CO)NH, HNCACB, HBHA(CBCACO)NH, HBHA(CBCA)NH, HNCO, HN(CA)CO, H(C)(CO)NH-TOCSY, WATERGATE		[61]
β-LG	DQF-COSY, TOCSY, NOESY	Thermal analysis	[62]
β-LG	^{13}C, ^{15}N-protein labelling 3D: ^{1}H-^{15}N NOESY-HSQC, ^{13}C NOESY-HSQC, HNHA, HNHB	Structure calculation (DYANA, X-PLOR)	[63]
β-LG	DQF-COSY, TOCSY, T1 and T2 relaxation studies	X-ray crystallography for modelling	[64]
β-LG	^{15}N,^{13}C-labelled proteins 3D: ^{13}C- and ^{15}N-edited NOESY-HSQC, HNHA, T1 and T2 ^{15}N relaxation times, WATERGATE	Structure calculation (ARIA extension of X-PLOR)	[4]
β-LG	^{1}H-^{15}N NOESY-HSQC, ^{13}C NOESY-HSQC		[65]
β-LG	TOCSY, NOESY, WATERGATE		[66]
β-LG	^{15}N, ^{13}C double-labelled protein 3D: CBCA(CO)NH, HNCACB, HNCO, HNCACO, relaxation analysis	CD	[67]
β-LG	^{15}N, ^{13}C double-labelled protein H/D Exchange experiments 3D: CBCA(CO)NH, HNCACB, HNCO, HNCACO	PCA	[68]
β-LG	^{1}H–^{15}N HSQC; H/D exchange; transverse relaxation (R2) 3D: HNCACB, CBCACONH, HNCO, HNCACO		[69]

3.1.2. β-Lactoglobulin

Bovine β-LG, due to its availability and great abundance, presents a remarkable model for molecular studies [2,70–73] and has been the most extensively studied milk protein for structural modelling using NMR. At pH below 3, the naturally existing dimeric form of β-LG dissociates into a monomeric form, which can retain the native conformation even at a pH value as low as 2.26 [71]. The first NMR resonance assignment of β-LG was performed by Molinari et al. [17], who revealed a highly structured β-sheet core for the monomeric form of β-LG at pH 2. At this pH, the monomer is unfolded, forming a β-sheet and random coils observed from numerous overlapping peaks. A similar study at pH 2 was published by Ragona et al. [18], who detected a β-core and 11 hydrophobic residues located around Trp_{19} and facing towards the interior of the monomer. Both studies used 1D and 2D NMR experiments (DQF-COSY, TOCSY and NOESY) without any isotopic labelling. Uhrinová et al. [63] for the first time proposed a completely resolved structure of a recombinant form of variant A of β-LG using a solution state of NMR. The study was performed using ^{13}C, ^{15}N-labelled proteins and different 3D heteronuclear pulse selective NMR techniques. The findings reported eight antiparallel β-sheets organized in a barrel and surrounded by α-helix. This study served as basic knowledge for further studies, which involved characterization of the structure and dynamics of β-LG. The coordinates of this study are available under code 1DV9 in PDB.

Another important study that observed the structure and kinetics of the folding of β-LG was published by Kuwata et al. [61], who also studied the monomeric state of the protein at pH 2. Their findings regarding the monomeric structure of β-LG were similar to those by Uhrinová et al. [63] with 8 antiparallel barrel-oriented β-sheets and one major α-helix; however, Kuwata et al. [61] also pointed out a Tanford transition of β-LG confirming rigid behavior of three β-sheets located under the α-helix and prompt fluctuation of the N and C terminals of the molecule. Kuwata et al. [61] performed their study using ^{13}C,^{15}N labelled β-LG and heteronuclear 2D and 3D NMR experiments. Structure calculation was completed using 1143 distance restraints for φ angle based on HNHA, χ_1 angle based on HNHB data and hydrogen bonds based on NOE data and solvent exchange [61]. The coordinates

from this study are available at PDB under the code 1cj5. The Tanford transition of β-LG, which results from the pH variation, was later studied in detail using heteronuclear NMR spectroscopy by Sakurai et al. [64]. Complete analysis of the NMR studies related to structural changes of β-LG are listed in Table 1.

3.2. Caseins

Contrary to major whey proteins, caseins as the main group of milk proteins have been less studied by NMR. Several studies observed the secondary structure orientation of defined regions of the polypeptide chains of individual caseins. This refers to the phosphorylated regions of the molecule of multi-phosphorylated motif Ser(P)-Ser(P)-Ser(P)-Glu-Glu, which has been shown to be critical for the interaction with amorphous calcium phosphate [74]. The peptide sequence with the Ser(P) residues that are of great importance in stabilization of casein phosphate complexes are the following [29]:

αs_1-CN (59-79): -Gln^{59}-Met-Glu-Ala-Glu-Ser(P)-Ile-Ser(P)-Ser(P)-Ser(P)-Glu-Glu-Ile-Val-Pro-Asn-Ser(P)-Val-Glu-Gln-Lys^{79}-

β-CN(1-25): -Arg^{1}-Glu-Leu-Glu-Glu-Leu-Asn-Val-Pro-Gly-Glu-Ile-Val-Glu-Ser(P)-Leu-Ser(P)-Ser(P)-Ser(P)-Glu-Glu-Ser-Ile-Thr-Arg^{25}-

αs_2-CN(1-21): -Lys^{1}-Asn-Thr-Met-Glu-His-Val-Ser(P)-Ser(P)-Ser(P)-Glu-Glu-Ser-Ile-Ile-Ser(P)-Gln-Glu-Thr-Tyr-Lys^{21}-

αs_2-CN(46-70): -Asn^{46}-Ala-Asn-Glu-Glu-Glu-Tyr-Ser-Ile-Gly-Ser(P)-Ser(P)-Ser(P)-Glu-Glu-Ser(P)-Ala-Glu-Val-Ala-Thr-Glu-Glu-Val-Lys^{70}-

These sequences have been studied extensively using the NMR for secondary structure identification, binding preferences for cations (Ca^{2+}) and aggregation preferences [25–33,75,76]. Most of the studies were focused on amino acids sequence analysis of αs_1-CN and β-CN, as the predominant caseins in milk and are thus important building blocks of the casein micelle. Complete assignment of the individual caseins using NMR methods has not been published so far. The following section summarizes the main studies that involve solution NMR for structural assignments of the four caseins (αs_1-, αs_2-, β- and κ-casein). Some of the studies involving structural elucidation using NMR are listed in the Table 2. The current structural studies performed using NMR of individual casein fragments are discussed in the following part of this review.

Table 2. NMR structural studies on peptides from caseins (β-, αs_1-, αs_2-, κ-casein) including the performed NMR methods and additional methods used. Abbreviations: WET - water suppression enhanced through T1 effect; sLED - suppression longitudinal encode decode. For other abbreviations see Table 1.

Protein Fragments	NMR Methods	Additional Methods	References
β-CN f(1–25)	2D: DQF-COSY, TOCSY, ROESY, NOESY		[25]
β-CN f(1–25)	2D: COSY, R-COSY, TOCSY, NOESY		[26]
β-CN f(1–25)	2D: DQF-COSY, TOCSY, NOESY, ROESY		[27]
β-CN f(1–25)	2D: DQF-COSY, TOCSY, NOESY, ROESY	Molecular modelling	[29]
α_{s1}-CN f(59–79)	2D: DQF-COSY, TOCSY, NOESY, ROESY	Molecular modelling	[30,76]
α_{s2}-CN f(1–23)	2D: TOCSY, NOESY, HSQC	FTIR, CD, Molecular modelling	[32]
α_{s1}-CN f(59–79)	2D: DQF-COSY, TOCSY, NOESY, WET, Presaturation	sLED, X-ray scattering	[28]
α_{s1}-CN f(136–196)	2D: DQF-COSY, TOCSY, NOESY, Presaturation	Far-UV CD, FTIR	[77]
α_{s2}-CN f(2–20)	2D: DQF-COSY, TOCSY, NOESY, WET, Presaturation	Molecular modelling	[31]
κ-CN f(98–111)	2D: DQF-COSY, TOCSY, ROESY		[33]
κ-CN f(130–153)	2D: DQF-COSY, TOCSY, NOESY, ROESY, Presaturation	CD	[16]
κ-CN f(1–44)	2D: DQF-COSY, TOCSY, NOESY	CD, Structure calculation (X-PLOR)	[34]

3.2.1. β-Casein

β-CN is a calcium-sensitive protein and binding of calcium ions to phosphorylated residues (five SerP residues) lead to its precipitation [6]. B-CN was predicted to have no or little secondary structure including α-helix 7–20%, 15–33% β-sheet, 20–30% β-turns and 20–25% polyproline II structure [20,78–80]. β-CN is strongly amphipathic with hydrophilic N-terminus (1–40), moderate hydrophobicity in the middle section (41–135) and highly hydrophobic C-terminus (136–209) [6].

The N-terminus of the polypeptide chain of β-CN has been of great interest to NMR studies due to phosphoserine motif (Ser_{15}, Ser_{17}, Ser_{18}, Ser_{19}, Glu_{20}, Glu_{21}) located in this part of the molecule, which is an important center of phosphorylation. The N terminus of the β-CN (β-CN f(1–25) was studied using NMR by Tsuda et al. [25] and Wahlgren et al. [26,75], who all used a purified peptide obtained after tryptic hydrolysis of β-CN. The NMR experiments were performed in H_2O/D_2O (90:10) and presence of calcium [25,26,75]. The main NMR methods used for these studies were 2D, including COSY, TOCSY and NOESY with adjusted phase and pulse sequence and water suppression. However, neither of these studies had observed existence of a specific structure on the observed peptide. In addition, Tsuda et al. [25] did not present any evidence for existence of β-turns in the peptide. From the NOE sequential assignments performed by Wahlgren et al. [26,75] it has been established that the N-terminus of the β-CN has multiple conformations with great flexibility.

Cross et al. [27] reported structural features of the peptide β-CN f(1–25) in the presence of cations, whereas a detailed structural model of β-CN f(1–25) complexed with amorphous calcium phosphate (ACP) was reported by Cross et al. [29]. Cross et al. [27] focused on the complete sequence specific assignment of the peptide in the presence of ammonium, sodium and calcium ions and of the dephosphorylated peptide with sodium ions. All the samples were prepared using 10 mM of peptide and specific preparation protocol depending of the ions and method used [27]. The existence of a cation-dependent specific structure was observed in sequential and non-sequential NOE connectivity in the amide (fingerprint), and the Hα region of the NOESY/ROESY spectra. Moreover, four structures were identified in β-CN peptide (1–25) in the presence of calcium, including one loop structure (Arg_1-Gly_4) and three β-turns (Val_8-Glu_{11}, $SerP_{17}$-Glu_{20} and Glu_{21}-Thr_{24}) [27].

The aforementioned findings by Cross et al. [27] were extended by Cross et al. [29] in further conformational studies of the same peptide (β-CN f(1–25)) complexed with ACP, giving the importance of calcium phosphate in the formation of the nanoclusters in the casein micelle. For the purpose casein phosphopeptide-ACP complexes were prepared as described by Reynolds [81] with a peptide concentration of 1.0–4.5 mM. NMR experiments included 2D methods (DQF-COSY, TOCSY and NOESY) using the States–Habercorn method [82] and the standard sequential assignment protocol [35,83]. The sequential assignment was completed in the fingerprint region of the NOESY spectra (NOEs and $^3J_{HNH\alpha}$ coupling constant) with mixing time of 250 ms where all residues of the complex were assigned. The proposed molecular model of CPP-ACP complex predicted an existence of a "patch" on the surface of the model consisting of Pro_9, Ile_{12}, Val_{13} and Glu_5 [29]. Moreover, it was proposed that the entire length of the peptide is involved in interactions with ACP due to many negatively charged residues. Hence, the Ser(P) motif does not appear to be the sole calcium binding motif for the complexation with calcium phosphate [29]. These findings using NMR experiments were a great contribution to the basic knowledge of the structural organization of the β-CN molecule. However, as previously discussed, the main focus to date has been on the N-terminus of the molecule and the sequential assignment of the whole molecule of β-CN has not been reported to date.

3.2.2. α_{s1}-Casein

α_{s1}-CN is also a calcium-sensitive casein with eight phosphorylated residues and a secondary structure consisting of, on average, 5–20% α-helix structure, 17–46% β-sheet and 29–35% β-turn structures [32,76,84]. This casein presents approximately 40% of total casein in bovine milk and is of great importance due to its association with calcium phosphate during the formation of the casein micelle. NMR was applied as a structural method, predominately for observations on parts of the

N-terminus, i.e., αs1-CN f(1–23) and f(59–79) [27,28,30–32] and the C-terminus, i.e., α_{s1}-CN f(136–196), as a region with strong affinity for self-association in the presence of salts [77].

α_{s1}-CN f(1–23) was studied by Malin et al. [32] and selected as an important part of the αs_1-CN molecule due to the high tendency for self-association [6]. In their study, Malin et al. [32] used NMR in combination with FTIR, CD and molecular modelling. The performed NMR experiments were homonuclear shift correlations including NOESY and TOCSY with mixing times 300 and 124 ms, respectively. For the heteronuclear experiments, HSQC was used for observing correlations among ^{1}H and ^{13}C nuclei. TOCSY spectra were used for observing amino acid patterns in the fingerprint region (NH-Nα), and HSQC was used for detection of cross peaks in the alpha region (Cα-Hα) and for identification of the Pro peak position. NOESY was used for sequential short-range NOE connectivity in the fingerprint region for all the residues except prolines. The Pro residues were observed to be in trans position that may lead to lower conformational fluctuations. The main observed structure was the poly-L-proline II (PPII) structure which has an important role in casein–casein interactions [32].

Another fragment of α_{s1}-CN that was used for structural studies using NMR was the phosphoserine region α_{s1}-CN f(59–79). Huq et al. [30] published complete sequence assignments of αs_1-CN fragment (59–79) in the presence of five moles of calcium per mol peptide and pH values of 4.06, 6.08 and 7.45. The NMR methods used were DQF-COSY, TOCSY and NOESY using the States–time-proportional phase incrementation (TPPI) method [85] for a phase sensitive mode. A standard procedure was used for a resonance assignment [35,83]. The intra-residue cross-peaks in the fingerprint region of the NOESY spectra were recorded in the best sequential assignments for spectra obtained at pH 6.08 [30]. The spectra recorded at other pH values produced weak and only a few non-sequential and sequential NOEs. The structural features observed in this work were Type-I β-turn involving residues 73–74 observed in the strong d_{NN} connectivity [30]. Hence, for Type-I β-turns a value of 2.6 Å can be expected which was observed in the d_{NN} (i + 2, i + 3) NOE connectivity [35]. Another region of the peptide, in which a specific conformation appeared, was residues 61–63, where a loop structure was observed [30]. Later, the obtained NOE constraints from this work were used in a molecular modelling study reported by the same authors [31]. It was confirmed that the turns and loops in the peptide have high a degree of flexibility and mobility, which promotes availability for interactions with calcium phosphate [31].

The complexation of α_{s1}-CN f(59–79) with calcium, fluoride and phosphate was studied by Cross et al. [28]. The NMR spectra of the complex were recorded as in the previous study [30], predominately recording DQF-COSY, TOCSY and NOESY spectra. The sequential assignment of the spin system was performed in the fingerprint region of NOESY spectra [28]. The medium range $d_{\gamma N}$(i, i + 2) in the region 72–74 and medium range NOE (i, i + 2) in the region assignments of $d_{N\alpha}$ 64–66 and $d_{N\alpha}$ 66–68 were implicated as β-turn conformations. In addition, NMR diffusion studies were used to record the radius of the core particle of the peptide complex with calcium, fluoride and phosphate resulting in a radius of approximately 2.12 nm [29]. From this work, Cross et al. [29] concluded that the secondary structure of the peptide, with or without complexing with calcium, fluoride or phosphate, reveals a similar conformation.

α_{s1}-CN f(136–196) was studied by Alaimo et al. [77], who dissolved the peptide in buffered solution of 90% H_2O/10% D_2O containing 10 mM Na_2HPO_4 and studied for structural characterization using NMR, FTIR and CD at variable temperature (10–70 °C). The NMR method selection was similar as in the previously discussed studies including DQF-COSY, TOCSY and NOESY and data acquisition using the TPPI method, as described by Marion et al. [85]. The results from Alaimo et al. [77] showed increased side chain mobility as temperature increased leading to a decreased amount of extended structures. However, at temperatures as high as 70 °C, the β-turn structures and some aromatic residues retain a stable position in the peptide chain. This was related to the involvement of Pro residues in hydrophobic turns for self-association of the peptide fragment which was referred to as a heat stable "molten globule" structural center in the α_{s1}-CN [77].

3.2.3. α_{s2}-Casein

αs_2-CN, due to high phosphorylation (10–13 SerP residues per molecule), is the most calcium sensitive protein [6]. However, since this casein is present only in 10% of the total caseins in milk and is comparatively difficult to isolate, only a few studies involved identification of the structural features of α_{s2}-casein using NMR. In αs_2-CN, two centers of phosphorylation have been observed, including 8–16 and 56–63. As in the previous studies on caseins, the NMR was used to study structural features of a defined region of the polypeptide chain. Huq et al. [31] published sequence specific assignments of the peptide αs_2-CN f(2–20) isolated from the parent molecule by tryptic digestion using calcium and ethanol and further purified by fast protein liquid chromatography (FPLC) and high performance liquid chromatography (HPLC). The NMR methods used were TOCSY, NOESY and DQF-COSY, recorded at pH 6.36 and temperature of −5 °C. The amino acid spin system was assigned in the fingerprint region by sequential and non-sequential NOE [35]. The study by Huq et al. [31] for the first time provided evidence of the tendency of the phosphoserine region of the peptide to form helical structure. The strong d_{NN} and weak $d_{\alpha N}$ NOE connectivity observed in the NOESY spectra proved existence of the helical structure in αs_2-CN f(2–20) [31].

Another peptide of αs_2-CN, f(46–70), was studied by Cross et al. [86]. They reported an NMR study of the peptide using the same methods and procedures as in the study of Huq et al. [3] and concluded that the phosphoserine motive of the peptide (46–70) reveals a similar amide chemical shift as observed in peptides β-CN (1–25), αs_1-CN (59–79) and αs_2-CN (2–20). However, the addition of calcium promoted distinctly variable conformations observed in the medium-range NOE connectivity in these four peptides [86].

3.2.4. κ-Casein

κ-CN is known to exist on the surface of the micelle where its hydrophilic C terminus (106–169) provides steric stabilization to the micelle [1]. Several studies that include the elucidation of the structure of peptides from the κ-CN molecule using NMR methods have been reported in the literature. Plowman et al. [33] presented complete chemical assignment of the peptide κ-CN f(98–111) in an attempt to predict the secondary structure. In this study, 20 mM of the peptide was suspended in DMSO-d_6 and studied at pH 3.0 and 7.5 and temperatures in the range 22–58 °C. The NMR methods used were DQF-COSY, TOCSY and ROESY performed in the phase-sensitive mode with the TPPI method, as described by Marion et al. [85]. Sequential and intra residue connectivities were accomplished using the standard assignment method described by Wüthrich [35]. The study by Plowman et al. [33] revealed the existence of extended or random structure of the peptide in solution with trans isomerism of the proline residues. Moreover, a pH increase from 3.0 to 7.5 resulted in a shifting of His residues downfield as a result of the deprotonation of the side chain protons of the amino acid. The temperature changes from 22 to 58 °C resulted only in small variations in the hydrogen bonding of the NH protons of His and Leu [33].

A larger portion of the N-terminus of κ-CN, κ-CN f(1–44) was studied by Bansal et al. [34]. In this study, 1 mM of peptide was dissolved in H_2O/D_2O/TFE (60/10/30) or D_2O/TFE (70/30). The standard NMR methods for structure observation were selected, including DQF-COSY for through bond coupling ($^3J_{HN\text{-}\alpha H}$), TOCSY for amino acids pattern and NOESY for structure calculation using NOE distance and dihedral angle restraints. The structure calculation was based on 11 dihedral angle restraints and 375 NOE restraints including intra residue, sequential, medium and long-range. This study, for the first time, reported the presence of a defined helix between Ile_{28} and Arg_{34} and an irregular helix between Ile_9 and Pro_{27} in the N terminal of κ-CN [34]. This was the first study that involved a structural observation of the casein peptide with a significant length using NMR.

For studying k-CN f(130–153), Plowman et al. [16] synthetized this peptide and dissolved it in H_2O/D_2O (90/10, v/v) or ^2H labelled TFE (trifluoroethanol). Similar to the previously discussed studies, the NMR methods used were DQF-COSY, TOCSY, ROESY, NOESY as the basic NMR experiment for sequential assignment of the polypeptide chain. The peptide dissolved in TFE produced a better

NOE peak dispersion in the NOESY and ROESY spectra in comparison to the peptide in the aqueous solution [16]. In this study, several conformational states were observed; the amino acids from Ile_{136} to Ser_{149} existed in α-helix conformation, Thr_{131}-Thr_{135} showed an unordered conformation, Glu_{151}-Ile_{153} extended structures and Thr_{145}-Ala_{148} a 3_{10} a helix. The Pro residues existed predominately in trans conformation [16].

4. NMR Studies on Casein Micelles

In addition to studies on specific segments on caseins, NMR has also been successfully applied to detect the position and interactions between colloidal calcium phosphate (CCP) nanoclusters and phosphoserine regions in caseins including liquid-state and solid-state ^{31}P and ^{43}Ca NMR [87–89]. The stability of CCP and its mobility was observed by Gonzalez-Jordan et al. [90]. For this purpose, they used magic-angle spinning (MAS) NMR acquiring ^{1}H-^{31}P cross-polarization spectra to detect the signal of immobile phosphorus. The signal for mobile phosphorus (organic and inorganic) was obtained by subtraction of the ^{31}P immobile signal after integration and deconvolution. Gonzales-Jordan et al. [90] reported that 81% of organic phosphate and 97% inorganic phosphate found in the micelle are in an immobile state when milk is in its native pH (6.7). Lowering of pH to an isoelectric point leads to increased protonation of both forms of phosphate leading to precipitation of caseins. pH recovering induced reorganization of the calcium phosphate nanoclusters giving similar values of mobility as under the native conditions of milk [90]. Thus, this finding offers a crucial benefit for understanding the changes in salt equilibria in casein micelles as a result of pH cycling that is used in the dairy industry for various processes.

Not only pH, but also temperature strongly affects the interactions among caseins and with CCP. Removal of β-casein, which is observed at low temperatures, was related to drastic changes of SerP structure and thus changes in the hydrophobicity and interaction preferences observed using liquid NMR [91]. The "loose" structure of the micelle and presence of internal passages offer advantages, such as liberation of the β-casein at low temperatures. Rollema and Brinkhuis [92] used NMR to observe casein micelle behavior as impacted by temperature (60–98 °C) and calcium removal. NMR was used in the aliphatic region of the NMR spectrum using the methyl resonances of Val, Ile and Leu, and aromatic regions for aromatic amino acids. It was observed that the caseins were characterized by great mobility, dissociated from the micelle and changed their conformational preferences at elevated temperature and calcium depletion. The changes were observed to be reversible in absence of whey proteins [92]. Combination of ^{31}P NMR and FTIR can also be used to observe the influence of ionic components in milk during temperature changes [93]. Moreover, using the chemical shifts assignments of ^{31}P NMR, FTIR can depict a qualitative allocation and transformation of different types of phosphorus in the milk presenting individual peaks, which are important for understanding mineral changes in casein micelles [93].

Solid-state NMR was also used to detect the presence of hydrogen bonding from non-phosphorylated parts of the caseins to amorphous calcium phosphate nanoclusters [94]. The protons of the side chains of Arg, Lys, Glu or Asp can be close in space with P atoms, with a distance of 3.4–4.4 Å, and hence the terminal groups of the side chains including R-COOH and R-NH_2 have direct interactions with the inorganic phosphate of CCP [94]. Similarly, Cross et al. [29], observing the β-casein peptide (1–25) interactions with amorphous calcium phosphate using NMR and molecular modeling, suggested that the entire length of the peptide is involved in interactions with calcium phosphate clusters. This refers predominately to negatively charged amino acids including Glu and SerP that are not included in the phosphoserine motifs. Moreover, these interactions mainly occur in hydrophilic regions of the molecules that are responsible for post-transitional changes of caseins [29].

Hindmarsh and Watkinson [95] confirmed the existence of phosphorus–calcium bonding in the micelle, which is not yet classified in literature. Using ^{1}H–^{31}P CP-MAS NMR in Mozzarella cheese and EDTA-chelated micelles, they observed the presence of immobile phosphates complexed with calcium which are not part of the binding between phosphoserine and calcium phosphate nanoclusters. In

addition, the proposed interactions included calcium linkages between individual phosphoserines in the casein micelle which can be within the same or with different caseins. However, the affinity of phosphoserines to bind calcium cation depends on their pKa which vary for individual amino acid in the polypeptide chain. The cation binding is greatest at the highest pKa. Phosphoserine centers in the caseins have a high pKa compared to individual phosphoserines in the polypeptide chain and thus are in the primary position for binding calcium ions [96]. Upon reaching saturation, other low pKa phosphoserines start to be involved in forming calcium bridges. This relationship was also confirmed using ^{31}P NMR chemical shifts in caseins in the presence of calcium cations [96]. Further saturation of the casein micelles with phosphate leads to structural changes in the micelle by formation of complexes or gelation by extending or compressing of the peptide chain. This was confirmed by FTIR and ^{31}P NMR on casein micelles with different phosphate content [97].

However, calcium binding to caseins balances hydrophobic interactions and electrostatic repulsions leading crosslinking and/or self-association of caseins [98]. Hydrophobic regions of caseins can form defined structures using hydrophobic interactions. Why then are they referred to as proteins with a flexible and undefined structure? The answer will likely be in the existence of high concentration of prolines in the peptide chain, which hinders the formation of defined structures and, as mentioned before, the existence of regions with hydrophilic nature. However, FTIR showed that caseins adapt certain structural features [20,32]. Hence, using FTIR and CD was confirmed for β-CN to have several defined structural regions in the molecules including helix, loops and polyproline II [20]. This was related to self-association and folding of the molecule in its hydrophobic regions. However, these methods cannot guarantee the exact position of amino acids linking. On the other hand, Malin et al. [32] included NMR in combination with FTIR, CD and molecular modeling of the N-terminal region of as_1-CN and identified the position of individual protons of amino acids. Thus, using methods for through-bond coupling they confirmed that all prolines are in trans conformation and with through space coupling they traced sequential $d_{\alpha N}(i, i + j)$ backbone assignments. Moreover, close positioning of amino acids leads to non-specific interactions including Van der Waals forces that allow for packing of peptide into irregular conformations [29].

5. Conclusions and Future Perspectives

NMR has been shown to be reliable and the most accurate method to observe the position of individual protons in the amino acid sequence, and thus, to identify the type of peptide connections in the protein when self-association takes place or other cross connections are established with the neighboring peptides. Thus, using multidimensional NMR methods through bond and space interactions, the complete sequential assignment and spatial orientation of a polypeptide chain can be achieved. The structure of α-LA and β-LG has been successfully observed and described using NMR and complete chemical shift assignment and tertiary structure can be found in the protein data bank. They have been fully described by multiple NMR methods, including three dimensional studies and isotopic labeling. Their globular structure and relatively short polypeptide chain have made the whey proteins of great interest to scientists as model proteins.

On the other hand, caseins, due to their conformational flexibility, have only been studied using two-dimensional NMR methods applied on specific polypeptide fragments of the molecule. Current resolved structures reported in literature using NMR proved that this method is highly reliable and can be used widely for protein studies, including caseins and casein micelle. The application of this method will enable innovative research and faster problem solutions for the industry as our understanding of the conformational behavior of these important proteins, under various relevant conditions, is unraveled.

In the dairy industry, there is a vast product evolution, generating new (and more complex) products where understanding and control of the protein structure are crucial to achieve the required structure, texture and stability. This leads to the need for using more sophisticated and powerful techniques for quality and process control. This review has demonstrated that NMR can provide

knowledge of the molecular level of mechanisms in milk proteins. Since NMR has proven to be a unique tool with high sensitivity to structural changes in milk proteins, its use may become more widespread, beyond academic research. Low-resolution NMR techniques are currently applied industrially for determining, e.g., solid fat content, but with advances in the technique and improved ease of use, it may also become applicable in the future for protein characterization in dairy products and ingredients.

Author Contributions: Authors T.M., T.V. and T.H. contributed equally to the writing, reviewing and revision of this review article. All authors have read and agreed to the published version of the manuscript.

Funding: The first author is funded by a Victoria University Research Training Stipend Scholarship (Grant number—INT010).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dalgleish, D.G.; Corredig, M. The structure of the casein micelle of milk and its changes during processing. *Annu. Rev. Food Sci. Technol.* **2012**, *3*, 449–467. [CrossRef]
2. Sawyer, L.; Holt, C. The secondary structure of milk proteins and their biological function. *J. Dairy Sci.* **1993**, *76*, 3062–3078. [CrossRef]
3. Forge, V.; Wijesinha, R.T.; Balbach, J.; Brew, K.; Robinson, C.V.; Redfield, C.; Dobson, C.M. Rapid collapse and slow structural reorganisation during the refolding of bovine α-lactalbumin. *J. Mol. Biol.* **1999**, *288*, 673–688. [CrossRef]
4. Uhrinová, S.; Smith, M.H.; Jameson, G.B.; Uhrin, D.; Sawyer, L.; Barlow, P.N. Structural changes accompanying pH-induced dissociation of the β-lactoglobulin dimer. *Biochemistry* **2000**, *39*, 3565–3574. [CrossRef] [PubMed]
5. Sakurai, K.; Konuma, T.; Yagi, M.; Goto, Y. Structural dynamics and folding of β-lactoglobulin probed by heteronuclear NMR. *Biochim. Et Biophys. Acta (BBA)-Gen. Subj.* **2009**, *1790*, 527–537. [CrossRef] [PubMed]
6. Huppertz, T. Chemistry of the caseins. In *Advanced Dairy Chemistry*; Springer: Boston, MA, USA, 2013; pp. 135–160.
7. Douglas, D.G. On the structural models of bovine casein micelles—Review and possible improvements. *Soft Matter* **2011**, *7*, 2265–2272.
8. De Kruif, C.G.; Huppertz, T.; Urban, V.S.; Petukhov, A.V. Casein micelles and their internal structure. *Adv. Colloid Interface Sci.* **2012**, *171*, 36–52. [CrossRef] [PubMed]
9. Huppertz, T.; Gazi, I.; Luyten, H.; Nieuwenhuijse, H.; Alting, A.; Schokker, E. Hydration of casein micelles and caseinates: Implications for casein micelle structure. *Int. Dairy J.* **2017**, *74*, 1–11. [CrossRef]
10. Horne, D.S. Casein micelle structure and stability. In *Milk Proteins*; Academic Press: Cambridge, MA, USA, 2020; pp. 213–250.
11. Gaucheron, F. The minerals of milk. *Reprod. Nutr. Dev.* **2005**, *45*, 473–483. [CrossRef]
12. Creamer, K.L.; Berry, G.P.; Mills, O.E. Study of the dissociation of beta casein from the bovine casein micelle at low temperature. *N. Z. J. Dairy Sci. Technol.* **1977**, *74*, 58–66.
13. Hyslop, D.B. Enzymatic coagulation of milk. In *Advanced Dairy Chemistry—1 Proteins*; Springer: Boston, MA, USA, 2003; pp. 839–878.
14. Anema, S.G.; Li, Y. Association of denatured whey proteins with casein micelles in heated reconstituted skim milk and its effect on casein micelle size. *J. Dairy Res.* **2003**, *70*, 73–83. [CrossRef] [PubMed]
15. Dupont, D.T.; Brodkorb, C.A.; Kouaouci, R. Quantitation of proteins in milk and milk products. In *Advanced Dairy Chemistry*; Springer: Boston, MA, USA, 2013; pp. 87–134.
16. Plowman, J.E.; Creamer, L.K.; Liddell, M.J.; Cross, J.J. Solution conformation of a peptide corresponding to bovine κ-casein B residues 130–153 by circular dichroism spectroscopy and 1 H-nuclear magnetic resonance spectroscopy. *J. Dairy Res.* **1997**, *64*, 377–397. [CrossRef] [PubMed]
17. Molinari, H.; Ragona, L.; Varani, L.; Musco, G.; Consonni, R.; Zetta, L.; Monaco, H.L. Partially folded structure of monomeric bovine β-lactoglobulin. *FEBS Lett.* **1996**, *381*, 237–243. [CrossRef]
18. Ragona, L.; Pusterla, F.; Zetta, L.; Monaco, H.L.; Molinari, H. Identification of a conserved hydrophobic cluster in partially folded bovine β-lactoglobulin at pH 2. *Fold. Des.* **1997**, *2*, 281–290. [CrossRef]
19. Kelly, S.M.; Jess, T.J.; Price, N.C. How to study proteins by circular dichroism. *Biochim. Et Biophys. Acta (BBA)-Proteins Proteom.* **2005**, *1751*, 119–139. [CrossRef] [PubMed]

20. Farrell, H.M., Jr.; Wickham, E.D.; Unruh, J.J.; Qi, P.X.; Hoagland, P.D. Secondary structural studies of bovine caseins: Temperature dependence of β-casein structure as analyzed by circular dichroism and FTIR spectroscopy and correlation with micellization. *Food Hydrocoll.* **2001**, *15*, 341–354. [CrossRef]
21. Lefèvre, T.; Subirade, M. Structural and interaction properties of β-Lactoglobulin as studied by FTIR spectroscopy. *Int. J. Food Sci. Technol.* **1999**, *34*, 419–428. [CrossRef]
22. Grewal, M.K.; Huppertz, T.; Vasiljevic, T. FTIR fingerprinting of structural changes of milk proteins induced by heat treatment, deamidation and dephosphorylation. *Food Hydrocoll.* **2018**, *80*, 160–167. [CrossRef]
23. Brownlow, S.; Cabral, J.H.M.; Cooper, R.; Flower, D.R.; Yewdall, S.J.; Polikarpov, I.; North, A.C.T.; Sawyer, L. Bovine β-lactoglobulin at 1.8 Å resolution—Still an enigmatic lipocalin. *Structure* **1997**, *5*, 481–495. [CrossRef]
24. Farrell, H.M.; Brown, E.M.; Malin, E.L. Higher order structures of the caseins: A paradox? In *Advanced Dairy Chemistry*; Springer: Boston, MA, USA, 2013; pp. 161–184.
25. Sakae, T.; Niki, R.; Kuwata, T.; Tanaka, I.; Hikichi, K. 1H NMR study of casein phosphopeptide (1–25): Assignment and conformation. *Magn. Reson. Chem.* **1991**, *29*, 1097–1102.
26. Wahlgren, N.M.; Léonil, J.; Dejmek, P.; Drakenberg, T. Two-dimensional nuclear magnetic resonance study of the β-casein peptide 1–25: Resonance assignments and secondary structure. *Biochim. Et Biophys. Acta (BBA)-Protein Struct. Mol. Enzymol.* **1993**, *1202*, 121–128. [CrossRef]
27. Cross, K.J.; Huq, N.L.; Bicknell, W.; Reynolds, E.C. Cation-dependent structural features of β-casein-(1–25). *Biochem. J.* **2001**, *356*, 277–286. [PubMed]
28. Cross, K.J.; Huq, N.L.; Stanton, D.P.; Sum, M.; Reynolds, E.C. NMR studies of a novel calcium, phosphate and fluoride delivery vehicle-αS1-casein (59–79) by stabilized amorphous calcium fluoride phosphate nanocomplexes. *Biomaterials* **2004**, *25*, 5061–5069. [CrossRef] [PubMed]
29. Cross, K.J.; Huq, N.L.; Reynolds, E.C. Casein phosphopeptide–amorphous calcium phosphate nanocomplexes: A structural model. *Biochemistry* **2016**, *55*, 4316–4325. [CrossRef] [PubMed]
30. Laila, H.N.; Cross, K.J.; Reynolds, E.C. A 1H-NMR study of the casein phosphopeptide αs1-casein (59–79). *Biochim. Et Biophys. Acta (BBA)-Protein Struct. Mol. Enzymol.* **1995**, *1247*, 201–208. [CrossRef]
31. Huq, N.L.; Cross, K.J.; Reynolds, E.C. Nascent helix in the multiphosphorylated peptide αS2-casein (2–20). *J. Pept. Sci.* **2003**, *9*, 386–392. [CrossRef] [PubMed]
32. Malin, E.L.; Alaimo, M.H.; Brown, E.M.; Aramini, J.M.; Germann, M.W.; Farrell, H.M.; McSweeney, P.L.H.; Fox, P.F. Solution structures of casein peptides: NMR, FTIR, CD, and molecular modeling studies of αs1-casein, 1–23. *J. Protein Chem.* **2001**, *20*, 391–404. [CrossRef]
33. Plowman, J.E.; Smith, M.H.; Creamer, L.K.; Liddell, M.J.; Coddington, J.; Gibson, J.J.; Engelbretsen, D.R. Proton assignment and structural features of a peptide from the chymosin-sensitive region of bovine k-casein determined by 2D-NMR spectroscopy. *Magn. Reson. Chem.* **1994**, *32*, 458–464. [CrossRef]
34. Bansal, P.S.; Grieve, P.A.; Marschke, R.J.; Daly, N.L.; McGhie, E.; Craik, D.J.; Alewood, P.F. Chemical synthesis and structure elucidation of bovine κ-casein (1–44). *Biochem. Biophys. Res. Commun.* **2006**, *340*, 1098–1103. [CrossRef]
35. Wüthrich, K. NMR with proteins and nucleic acids. *Europhys. News* **1986**, *17*, 11–13. [CrossRef]
36. Kanelis, V.; Forman-Kay, J.D.; Kay, L.E. Multidimensional NMR methods for protein structure determination. *Iubmb Life* **2001**, *52*, 291–302. [CrossRef]
37. Tugarinov, V.; Kanelis, V.; Kay, L.E. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat. Protoc.* **2006**, *1*, 749–754. [CrossRef] [PubMed]
38. Wüthrich, K. Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem.* **1990**, *265*, 22059–22062. [PubMed]
39. Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G., III; Skelton, N.J. *Protein NMR Spectroscopy: Principles and Practice*; Elsevier: Amsterdam, The Netherlands, 1995.
40. Bodenhausen, G.; Ruben, D.J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **1980**, *69*, 185–189. [CrossRef]
41. Marion, D.; Wüthrich, K. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1 H-1 H spin-spin coupling constants in proteins. In *NMR in Structural Biology: A Collection of Papers by Kurt Wüthrich*; World Scientific: Singapore, 1995; pp. 114–121.
42. Griesinger, C.; Otting, G.; Wüthrich, K.; Ernst, R.R. Clean TOCSY for proton spin system identification in macromolecules. *J. Am. Chem. Soc.* **1988**, *110*, 7870–7872. [CrossRef]

43. Wüthrich, K.; Wider, G.; Wagner, G.; Braun, W. Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proteins nuclear magnetic resonance. *J. Mol. Biol.* **1982**, *155*, 311–319. [CrossRef]
44. Wüthrich, K.; Billeter, M.; Braun, W. Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton–proton distances. In *NMR in Structural Biology. A Collection of Papers by Kurt Wüthrich*; World Scientific: Singapore, 1995; pp. 218–243.
45. Karplus, M. Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.* **1959**, *30*, 11–15. [CrossRef]
46. Ramachandran, G.N. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **1963**, *7*, 95–99. [CrossRef]
47. Mariette, F. NMR relaxometry and imaging of dairy products. *Mod. Magn. Reson.* **2018**, 1535–1557. [CrossRef]
48. Brew, K. α-Lactalbumin. In *Advanced Dairy Chemistry—1 Proteins*; Springer: Boston, MA, USA, 2003; pp. 387–419.
49. Ikeguchi, M.; Kuwajima, K.; Mitani, M.; Sugai, S. Evidence for identity between the equilibrium unfolding intermediate and a transient folding intermediate: A comparative study of the folding reactions of alpha.-lactalbumin and lysozyme. *Biochemistry* **1986**, *25*, 6965–6972. [CrossRef]
50. Quezada, C.M.; Schulman, B.A.; Froggatt, J.J.; Dobson, C.M.; Redfield, C. Local and global cooperativity in the human α-lactalbumin molten globule. *J. Mol. Biol.* **2004**, *338*, 149–158. [CrossRef] [PubMed]
51. Schulman, B.A.; Kim, P.S.; Dobson, C.M.; Redfield, C. A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat. Struct. Biol.* **1997**, *4*, 630–634. [CrossRef] [PubMed]
52. Alexandrescu, A.T.; Evans, P.A.; Pitkeathly, M.; Baum, J.; Dobson, C.M. Structure and dynamics of the acid-denatured molten globule state of alpha.-lactalbumin: A two-dimensional NMR study. *Biochemistry* **1993**, *32*, 1707–1718. [CrossRef] [PubMed]
53. Wijesinha-Bettoni, R.; Dobson, C.M.; Redfield, C. Comparison of the structural and dynamical properties of holo and apo bovine α-lactalbumin by NMR spectroscopy. *J. Mol. Biol.* **2001**, *307*, 885–898. [CrossRef]
54. Redfield, C.; Schulman, B.A.; Milhollen, M.A.; Kim, P.S.; Dobson, C.M. α-Lactalbumin forms a compact molten globule in the absence of disulfide bonds. *Nat. Struct. Biol.* **1999**, *6*, 948–952. [CrossRef]
55. Ramboarina, S.; Redfield, C. Structural characterisation of the human α-lactalbumin molten globule at high temperature. *J. Mol. Biol.* **2003**, *330*, 1177–1188. [CrossRef]
56. Alexandrescu, A.T.; Broadhurst, R.W.; Wormald, C.; Chyan, C.; Baum, J.; Dobson, C.M. 1H-NMR assignments and local environments of aromatic residues in bovine, human and guinea pig variants of α-lactalbumin. *Eur. J. Biochem.* **1992**, *210*, 699–709. [CrossRef]
57. Berliner, L.J.; Kaptein, R.; Koga, K.; Musci, G. NMR Studies of the Structure and Environment of the Milk Protein α-Lactalbumin. In *NMR Applications in Biopolymers*; Springer: Boston, MA, USA, 1990; pp. 231–253.
58. Akio, S.; Ikeguchi, M.; Kobayashi, T.; Sugai, S. A synthetic peptide study on the molten globule of α-lactalbumin. *J. Biochem.* **1996**, *119*, 947–952.
59. Kazuo, K.; Hoshino, M.; Era, S.; Batt, C.A.; Goto, Y. α→ β transition of β-lactoglobulin as evidenced by heteronuclear NMR. *J. Mol. Biol.* **1998**, *283*, 731–739.
60. Fogolari, F.; Ragona, L.; Zetta, L.; Romagnoli, S.; de Kruif, G.K.; Molinari, H. Monomeric bovine β-lactoglobulin adopts a β-barrel fold at pH 2. *FEBS Lett.* **1998**, *436*, 149–154. [CrossRef]
61. Kazuo, K.; Era, S.; Hoshino, M.; Forge, V.; Goto, Y.; Batt, C.A. Solution structure and dynamics of bovine β-lactoglobulin A. *Protein Sci.* **1999**, *8*, 2541–2545.
62. Belloque, J.; Smith, G.M. Thermal denaturation of β-lactoglobulin. A 1H NMR study. *J. Agric. Food Chem.* **1998**, *46*, 1805–1813. [CrossRef]
63. Uhrinova, S.; Uhrin, D.; Denton, H.; Smith, M.; Sawyer, L.; Barlow, P.N. Complete assignment of 1 H, 13 C and 15 N chemical shifts for bovine β-lactoglobulin: Secondary structure and topology of the native state is retained in a partially unfolded form. *J. Biomol. NMR* **1998**, *12*, 89–107. [CrossRef] [PubMed]
64. Belloque, J.; López-Fandiño, R.; Smith, G.M. A 1H-NMR study on the effect of high pressures on β-lactoglobulin. *J. Agric. Food Chem.* **2000**, *48*, 3906–3912. [CrossRef]
65. Sakurai, K.; Goto, Y. Dynamics and mechanism of the Tanford transition of bovine β-lactoglobulin studied using heteronuclear NMR spectroscopy. *J. Mol. Biol.* **2006**, *356*, 483–496. [CrossRef]
66. Kazuo, K.; Li, H.; Yamada, H.; Batt, C.A.; Goto, Y.; Akasaka, K. High pressure NMR reveals a variety of fluctuating conformers in β-lactoglobulin. *J. Mol. Biol.* **2001**, *305*, 1073–1083.

67. Lübke, M.; Guichard, E.; Tromelin, A.; le Quéré, J.L. Nuclear magnetic resonance spectroscopic study of β-lactoglobulin interactions with two flavor compounds, γ-decalactone and β-ionone. *J. Agric. Food Chem.* **2002**, *50*, 7094–7099. [CrossRef]
68. Kazumasa, S.; Goto, Y. Principal component analysis of the pH-dependent conformational transitions of bovine β-lactoglobulin monitored by heteronuclear NMR. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15346–15351.
69. Kazumasa, S.; Yagi, M.; Konuma, T.; Takahashi, S.; Nishimura, C.; Goto, Y. Non-native α-helices in the initial folding intermediate facilitate the ordered assembly of the β-barrel in β-lactoglobulin. *Biochemistry* **2017**, *56*, 4799–4807.
70. Sawyer, L.; Brownlow, S.; Polikarpov, I.; Wu, S. β-Lactoglobulin: Structural studies, biological clues. *Int. Dairy J.* **1998**, *8*, 65–72. [CrossRef]
71. Sawyer, L. β-Lactoglobulin. In *Advanced Dairy Chemistry*; Springer: Boston, MA, USA, 2013; pp. 211–259.
72. Kinsella, J.E.; Whitehead, D.M. Proteins in whey: Chemical, physical, and functional properties. In *Advances in Food and Nutrition Research*; Academic Press: Cambridge, MA, USA, 1989; Volume 33, pp. 343–438.
73. Kontopidis, G.; Holt, C.; Sawyer, L. Invited review: β-lactoglobulin: Binding properties, structure, and function. *J. Dairy Sci.* **2004**, *87*, 785–796. [CrossRef]
74. Reynolds, E.C. Anticariogenic complexes of amorphous calcium phosphate stabilized by casein phosphopeptides: A review. *Spec. Care Dent.* **1998**, *18*, 8–16. [CrossRef] [PubMed]
75. Wahlgren, N.; Dejmek, M.P.; Drakenberg, T. Secondary structures in β-casein peptide 1–42: A two dimensional nuclear magnetic resonance study. *J. Dairy Res.* **1994**, *61*, 495–506. [CrossRef] [PubMed]
76. Huq, N.L.; Cross, K.J.; Reynolds, E.C. Molecular modelling of the multiphosphorylated casein phosphopeptide [alpha] S1-casein (59–79) based on NMR constraints. *J. Dairy Res.* **2004**, *71*, 28. [CrossRef] [PubMed]
77. Alaimo, M.H.; Farrell, H.M., Jr.; Germann, M.W. Conformational analysis of the hydrophobic peptide αs1-casein (136–196). *Biochim. Et Biophys. Acta (BBA)-Protein Struct. Mol. Enzymol.* **1999**, *1431*, 410–420. [CrossRef]
78. Creamer, L.K.; Richardson, T.; Parry, D.A.D. Secondary structure of bovine αs1-and β-casein in solution. *Arch. Biochem. Biophys.* **1981**, *211*, 689–696. [CrossRef]
79. Graham, E.R.B.; Malcolm, G.N.; McKenzie, H.A. On the isolation and conformation of bovine β-casein A1. *Int. J. Biol. Macromol.* **1984**, *6*, 155–161. [CrossRef]
80. Qi, P.X. Studies of casein micelle structure: The past and the present. *Le Lait* **2007**, *87*, 363–383. [CrossRef]
81. Reynolds, E.C. Stabilized Calcium Phosphate Complexes. U.S. Patent 8,609,071, 17 December 2013.
82. Haberkorn, R.A.; Ruben, D.J. A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.* **1982**, *48*, 286–292.
83. Chazin, W.J.; Wright, P.E. A modified strategy for identification of 1H spin systems in proteins. *Biopolym. Orig. Res. Biomol.* **1987**, *26*, 973–977.
84. Byler, D.M.; Farrell, H.M., Jr.; Susi, H. Raman spectroscopic study of casein structure. *J. Dairy Sci.* **1988**, *71*, 2622–2629. [CrossRef]
85. Dominique, M.; Ikura, M.; Tschudin, R.; Bax, A.D. Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **1989**, *85*, 393–399.
86. Cross, J.K.; Huq, N.L.; He, H.; Stanton, D.P.; Lau, K.; Reynolds, E.C. Structural characterization of the anticariogenic casein phosphopeptide αS2-casein (46–70) complexed with amorphous calcium phosphate. *Aust. Dent. J.* **2007**, *52*, S10–S11. [CrossRef]
87. Bak, M.; Rasmussen, L.K.; Petersen, T.E.; Nielsen, N.C. Colloidal calcium phosphates in casein micelles studied by slow-speed-spinning 31P magic angle spinning solid-state nuclear magnetic resonance. *J. Dairy Sci.* **2001**, *84*, 1310–1319. [CrossRef]
88. Belloque, J. High-resolution NMR of milk and milk proteins. In *Modern Magnetic Resonance*; Springer: Dordrecht, The Netherlands, 2008; pp. 1631–1635.
89. Wahlgren, N.M.; Dejmek, P.; Drakenberg, T. A 43 Ca and 31 P NMR study of the calcium and phosphate equilibria in heated milk solutions. *J. Dairy Res.* **1990**, *57*, 355–364. [CrossRef]
90. Gonzalez-Jordan, A.; Thomar, P.; Nicolai, T.; Dittmer, J. The effect of pH on the structure and phosphate mobility of casein micelles in aqueous solution. *Food Hydrocoll.* **2015**, *51*, 88–94. [CrossRef]
91. Ishii, T.; Hiramatsu, K.; Ohba, T.; Tsutsumi, A. The liquid-state 31P-nuclear magnetic resonance study on microfiltrated milk. *J. Dairy Sci.* **2001**, *84*, 2357–2363. [CrossRef]

92. Rollema, H.S.; Brinkhuis, J.A. A 1 H-NMR study of bovine casein micelles; influence of pH, temperature and calcium ions on micellar structure. *J. Dairy Res.* **1989**, *56*, 417–425. [CrossRef]
93. Boiani, M.; Fenelon, M.; FitzGerald, R.J.; Kelly, P.M. Use of 31P NMR and FTIR to investigate key milk mineral equilibria and their interactions with micellar casein during heat treatment. *Int. Dairy J.* **2018**, *81*, 12–18. [CrossRef]
94. De Sa Peixoto, P.; Silva, J.V.C.; Laurent, G.; Schmutz, M.; Thomas, D.; Bouchoux, A.; Gesan-Guiziou, G. How high concentrations of proteins stabilize the amorphous state of calcium orthophosphate: A solid-state nuclear magnetic resonance (NMR) study of the casein case. *Langmuir* **2017**, *33*, 1256–1264. [CrossRef]
95. Hindmarsh, J.P.; Watkinson, P. Experimental evidence for previously unclassified calcium phosphate structures in the casein micelle. *J. Dairy Sci.* **2017**, *100*, 6938–6948. [CrossRef] [PubMed]
96. Baumy, J.-J.; Guenot, P.; Sinbandhit, S.; Brulé, G. Study of calcium binding to phosphoserine residues of β-casein and its phosphopeptide (1–25) by 31 P NMR. *J. Dairy Res.* **1989**, *56*, 403–409. [CrossRef]
97. Van Hekken, D.L.; Dudley, R.L. Analysis of Modified Whole Casein with Different Phosphorus Contents Using Phosphorus-31 Nuclear Magnetic Resonance and Fourier Transform Infrared Spectroscopy. *J. Dairy Sci.* **1997**, *80*, 2751–2759. [CrossRef]
98. Swaisgood, H.E. Chemistry of the caseins. In *Advanced Dairy Chemistry—1 Proteins*; Springer: Boston, MA, USA, 2003; pp. 139–201.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

2C. The importance of Fourier Transform Infrared spectroscopy (FTIR) in structural studies of milk proteins

2C.1 Introduction

Fourier Transform Infrared spectroscopy (FTIR) is a spectroscopic tool that have been successfully used for identification and verification of chemical compounds. The use of FTIR in structural studies of milk proteins aims at establishing the relationship between the protein functionality and structure. The use of FTIR in food analysis has been available in research since early 1970s (Griffiths *et al*., 1983, Van de Voort, 1992). The FTIR spectroscopy can identify how different processing conditions on the milk products affects the milk proteins structures. Moreover, FTIR can successfully detect the heating induced structural changes of milk proteins in raw and concentrated skim milk (Grewal *et al*., 2018; Markoska *et al*., 2018; Mediwaththe *et al*., 2018; Markoska *et al*., 2019), the effect of deamidation and dephosphorylation on protein structure (Grewal *et al*., 2018), the structural changes of proteins in milk powders induced by the changes in the temperature (Gaiani *et al*., 2010; Kher *et al*., 2007; Haque *et al*., 2011; Zhou *et al*., 2006; Ye *et al*., 2017), characterizing the genetic protein variants between A1 and A2 milk (Daniloski *et al*., 2022b) and many more. The FTIR spectroscopy was used for identification of structural components of peptide fragments and opioid peptides form caseins (αs-CN and β-CN) (Malin *et al*., 2001; Markoska *et al*., 2021; Markoska *et al*., 2022). The FTIR provide inexpensive and rapid analysis, with minimum time for sample preparation. The coupling of FTIR with multivariate approaches leads to many advantages in improving the signal sensitivity due to better signal to noise ratio, capability to separate the samples into groupings according to selected variables and higher accuracy. In addition, to improve the quality of the spectra and separation and identification of peaks several methods can be applied on the original FTIR spectra. Some of the most used methods are deconvolution, peak fitting, derivatization, principal component analysis (PCA), 2D correlation and many more (Byler and Susi, 1986; Susi, and Byler, 1983; Noda and Ozaki, 2005).

2C.2 FTIR instrumentation

The FTIR spectrometer contain a source of infrared radiation, a process for resolving the infrared radiation into the wavelength and detector. In the FTIR instruments the wavelength arrive in the spectrometer simultaneously where the signal mathematically converts into typical IR spectrum (Rodriguez-Saona *et al*., 2017; Van de Voort, 1992). The instrument uses interferometer where IR from the source is split by a beam splitter into two beams. Each part of the beam goes to either moving or fixed mirror, then the reflected beams in the beam splitter are recombined resulting to interference which is pointed to the sample cell and then to the detector. The moving mirror results with changes of the optical path length of the both half beams producing interferogram. The interferogram consists three interference states or intermediate, constructive and destructive which intensity is measured by the detector. The collected data in the detector then is mathematically transformed or Fourier Transformed from time domain into frequency domain or IR spectrum. The example diagram of the interferometer is presented in Figure 1.

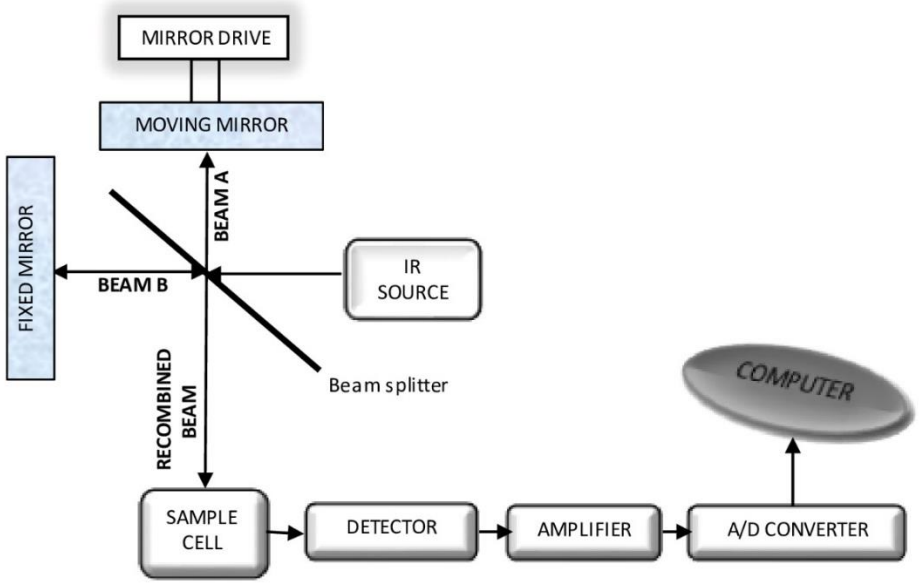


Figure 1. Diagram of interferometer

The infrared regions that are significant for food analysis are classified as near (NIR) and mid infrared region (MIR) which appear in regions of 12500-4000 cm^{-1} and 4000-600 cm^{-1}, respectively (Barth, 2007). NIR infrared spectra due to low structural sensitivity is less used for structural analysis (Karoui *et al*., 2010). However, MIR spectra can easily detect the conformational state of the organic samples including milk proteins and milk fats. The MIR region of the FTIR spectra extends in range of 4000 to 600 cm^{-1}. In the spectra the x-axis presents the wavenumbers or wavelengths and the y-axis presents either absorbance or transmittance. The vibrational frequencies appearing in the spectra result from the strength of the vibrating bonds and the mass of the atoms. The IR absorption frequencies are specific for different organic functional groups. The development of the attenuated total reflectance (ATR) technology brought great advantages in qualitative and quantitative FTIR analysis (Braue and Pannella, 1987). The ATR operates using infrared light, which is directed between the ATR crystal and the sample. The ATR requires minimum or no sample preparation provides consistent spectra collection, higher energy and improved resolution, better signal-to-noise ratio and eliminates the variations of the cell path lengths. This instrumentation is a key to simple and robust operation producing high quality data for food analysis.

2C.3 FTIR regions and spectra processing for protein studies

The FTIR analysis of food systems is characterised with specific absorption bands that are associated with individual components. The components that are generally observed are proteins, fats and carbohydrates that all contribute to the obtained spectrum (Van de Voort, 1992). In the FTIR spectrum, the infrared absorption of the proteins and peptides is presented in Table 1. Moreover, each region corresponds to different stretching intensity of the molecule. The regions are classified accordingly to the vibrational intensity and have been previously classified as A, B and Amide I-VII (Carbonaro and Nucara, 2010; Barth, 2007; Iñón *et al*., 2004; Zhou *et al*., 2006; Karoui *et al*., 2003; Grewal *et al*., 2017).

Table 1. FTIR regions for absorption of proteins and peptides

Spectra regions	Band Description	Wavenumbers (cm^{-1})
Amide A	N-H, O-H stretching	3700 – 3000 cm^{-1}
Amide B	N-H stretching; CH_2 asymmetric stretch	3000 – 2800 cm^{-1}
Amide I	C=O stretching	1700 – 1600 cm^{-1}
Amide II	C-N stretching, N-H bending	1600 – 1500 cm^{-1}
Amide III	C-N stretching; N-H bending	1500 – 1200 cm^{-1}
Amide IV	O-C-N bending	767 – 625 cm^{-1}
Amide V	Out-of-plane N-H bending	800 – 640 cm^{-1}
Amide VI	Out-of-plane C=O bending	600 – 500 cm^{-1}
Amide VII	Skeletal torsion	200 cm^{-1}

For structural studies of milk proteins, the most observed regions are Amide I, II and III. Moreover, these regions provide particular frequencies that assign the secondary structure of proteins. The average bands that are correlated to specific secondary structure from Amide I include bands at 1645 cm^{-1} (α-helix), 1633 and 1682 cm^{-1} (intramolecular/antipallarel β-sheets), 1625 cm^{-1}(intermolecular/ parallel β-sheets), 1673 cm^{-1} (turns) and 1645 cm^{-1} (random coil) (Grewal *et al*., 2017; Carbonaro and Nucara, 2010; Pelton and McLean, 2000). The absorption in Amide I regions is a result of the stretching of the peptide bond (C=O) of the protein. The Amide II regions provide bands with lower intensity in comparison to Amide I. The average bands that correspond to secondary structure of proteins and peptides include 1520 cm^{-1} (antiparallel β-sheets) and 1540 cm^{-1} (parallel β-sheets) (Pelton and McLean, 2000). In addition, in Amide II region the absorbance can also be related to the side chains of amino acids. Thus, the amino acids that contribute to the signal in this region are aspartic acid, asparagine, arginine, glutamine, glutamic acid, lysine and tyrosine (Mazerolles *et al*., 2001). The Amide III region shows a weak signal that also can assign to secondary structure components of

proteins. The observed frequencies in this region are for α-helix, β-sheets and unordered structures (1328 – 1289 cm^{-1}, 1255-1224 cm^{-1} and 1288-1256 cm^{-1}, respectively) (Grewal *et al*., 2017). The common problem that arise on FTIR analysis on milk is the water signal appearing at 1645 cm^{-1}, 2125 cm^{-1} and 3400 cm^{-1} (Yang *et al*., 2015). The water signal overlap with the absorption of the bands in Amide I region. To avoid the water interference the water signal can be easily removed by subtracting the blank spectrum from the sample spectrum. It is important that the blank spectrum be analysed under identical conditions as the sample spectrum. This includes temperature, pH, concentration, number of scans; cell path length and spectral resolution are kept identical for both blank and sample (Yang *et al*., 2015).

For better peak identification and peak resolution of the Amide I region the FTIR spectra is processed using different methods provided by the processing software. Some of the common methods for better peak identification include deconvolution, peak fitting, derivatization, principal component analysis (PCA), and 2D correlation to name a few (Byler and Susi, 1986, Susi and Byler, 1983; Noda and Ozaki, 2005). The deconvolution procedure alters the shape of the bands including anodization and line shaping function. This function is accompanied with curve fitting procedure, which provides better peak identification and resolving the overlapped bands into distinct peaks (Byler and Susi, 1986). A Lorentzian line approximates the intrinsic line shape of the peak.

$$E_{0(\nu)} = \frac{\sigma/\pi}{\sigma^2 + \nu^2} \quad [1]$$

Where $E_{0(\nu)}$ is the intrinsic line shape function of the spectrum $E(\nu)$ and σ is the line width.

The second derivative analysis of the FTIR spectra provides distinctive band frequencies that are characteristic for specific secondary structure. The bands in second derivative spectra define certain secondary structure component of proteins that was confirmed to be nearly identical with the information that was computed from crystallographic data (Dong, 1990). The shape of the infrared spectrum is calculated using Lorentzian function as shown in equation 2 (Susi and Byler, 1983) where the

absorbance is ***a, s*** is the width with half height and the frequency with reference to the line centre is ***v***:

$$A = (\frac{s}{\pi})/(s^2 + v^2) \quad [2]$$

The second derivative is calculated in equation 3 where $\boldsymbol{\alpha}$ is $1/s^2$ and the peak frequency is same with the original peak frequency. For the second derivative spectra, the peak intensity is proportional to the intensity of the original peak and inversely proportional to the square of the half width (Susi and Byler, 1983).

$$A^{"} = -(1/\pi s)[2a(1 - 3av^2)/(1 + av^2)^3] \quad [3]$$

The PCA is a bilinear modelling method that give information that is present in the multidimensional set. PCA provides information by generating a set of principal components (PCs) as coordinated axes with lowest possible loss of information. The groupings of different samples are presented in the score plots based on the variables, on the other hand the loading plots presents the wavenumbers with high loading which classified the samples into different groups. The aim of PCA analysis is to find the linear combinations of the main variables that differentiate the observed samples.

The 2D correlation analysis presents a stacked-trace or 3D representation of the spectrum providing best overall view of the intensity by providing sharper and better resolving peaks than the corresponding traditional 1D spectrum (Noda and Ozaki, 2005). The 2D correlation plot presents positive and negative cross peaks positioned at the off diagonal position of the plot confirming correlation features between the FTIR bands. The correlation can be obtained from the bands originating from the same spectroscopic data so called homonuclear correlation and correlation among the band from different types of spectroscopic data e.g. FTIR vs Raman. The advantage of the use of 2D correlation of the FTIR spectra arise from simplification of the overlapping peaks, improvement of the resolution, investigating the sequential order of spectra intensity produced from external perturbation (modified sample conditions), determining clear assignment using the correlation of bands and many more.

2C.4 FTIR application in milk protein structural studies

The use of FTIR in studying the conformational characteristics of milk proteins have been widely used (Byler and Susi, 1986; Kumosinski and Farrell Jr, 1993; Malin *et al*., 2001; Farrell Jr *et al*., 2001; Lefèvre and Subirade, 1999; Grewal *et al*., 2017; Markoska *et al*., 2019a; Markoska *et al*., 2019b; Daniloski *et al*., 2022). The use of FTIR spectrum for observation of the structure of milk proteins involve mainly two strong band absorption found in Amide I and II region. Thus, in Amide I region the strong band intensity is observed at 1650 cm^{-1} that appears from the C=O stretching of the peptide bond. In Amide II region, the strongest band absorption is at 1550 cm^{-1} that appear from N-H bending deformation with 60 % and C-N stretching vibrations that is 40 %. In Amide II region the band intensity is gradually, lower than the band intensity from Amide I region. The typical FTIR spectrum of milk protein in Amide I and II is shown in Figure 2.

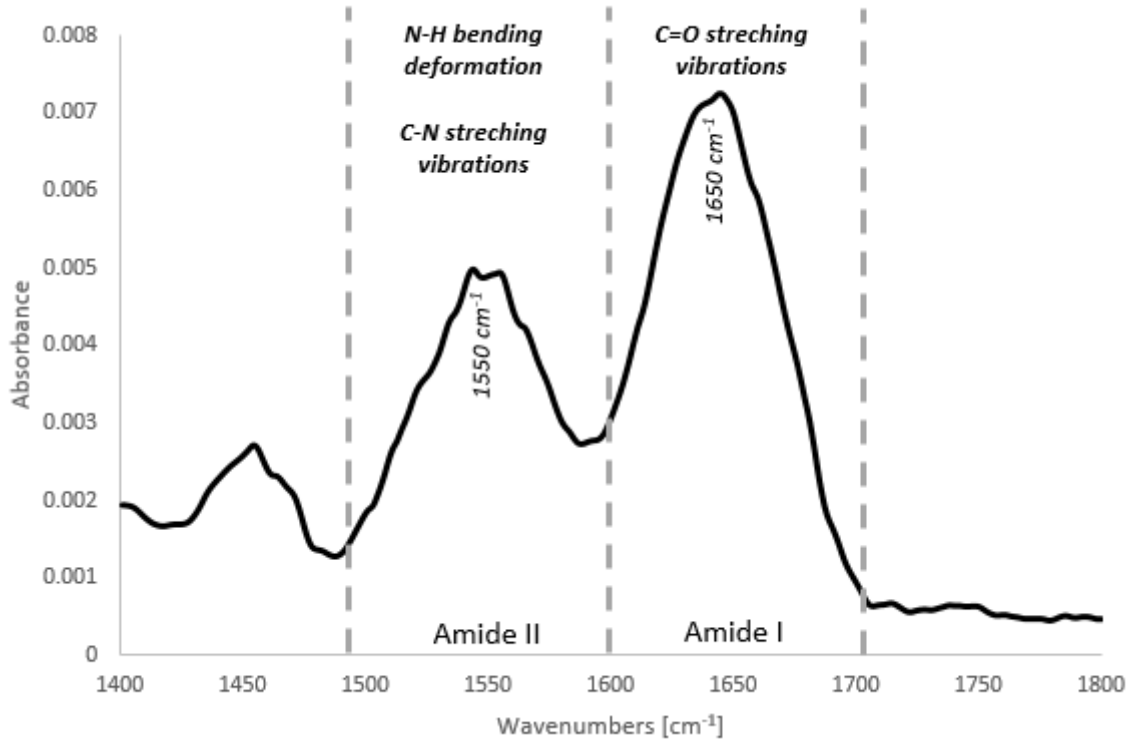


Figure 2. Typical FTIR spectra of milk protein in Amide I and II region.

The accuracy of the FTIR spectrum of milk proteins can be affected by many factors that can be classified as instrumental, physicochemical and biological (Dupont *et al*., 2013). The instrumental factors include the instrument conditions like instrument calibration, temperature, linearity, homogenizations to name a few. The physicochemical factors refer to the presence of carboxylic acid and the proportion of non-protein nitrogen in the observed sample. Moreover, the variation in the non-protein nitrogen affects the number of peptide bonds and thus can affect the accuracy of the instrument. The variation of the presence of carboxylic acid leads to interference absorption of the protein spectrum that usually occurs due to fermentation of lactose and greater presence of citrate in the milk samples (Sjaunja and Andersson, 1985). The biological factor that affects the accuracy of the FTIR spectrum include the breed, species, and stage of lactation, season and feeding. These factors can affect the accuracy of the FTIR spectra and whenever possible the instrument needs to be calibrated and adapted to the new conditions.

The secondary structure components of milk proteins were analysed extensively using FTIR spectroscopy in order to detect the structural variation in regards to the different processing conditions. The structural elements of milk proteins including whey proteins, caseins and peptides isolated from caseins identified by FTIR spectroscopy are discussed individually.

2C.4.1 FTIR studies on whey proteins

The employment of FTIR in structural analysis of milk proteins have been predominantly involved to determine the structural transition of the secondary structure components during processing. The structure of whey proteins have been intensively studied using this method to confirm the interaction preferences, aggregation, and dissociation.

β-Lactoglobulin (β-LG)

The structure of β-LG was successfully resolved using X-ray crystallography and NMR spectroscopy (Uhrínová *et al*., 1998; Molinari *et al*., 1996; Bewley *et al*., 1997; Qin *et al*., 1998) concluding with experimental values in solution of 8 % α-helix, 45 % β-sheets and 47 % random coils. The employment of FTIR analysis for observation of the structure of β-LG involved several studies that observed the structural transition of the protein concerning the observed conditions (Lefèvre and Subirade, 1999; Boye *et al*., 1996; QI *et al*., 1997; Bhattacharjee *et al*., 2005; Rahaman *et al*., 2015). The FTIR analysis of β-LG molecule confirmed presence of concentration, pH and thermal induced modification of the spectrum (Lefèvre and Subirade, 1999; QI *et al*., 1997). The FTIR spectrum successfully can identify the structural and interaction properties of the molecule by identifying the difference between the monomeric, polymeric and aggregated structural condition.

α-Lactalbumin(α-LA)

Prestrelski *et al*., (1991) performed the structural analysis of α-LA using FTIR identifying the two types of α-helix secondary structures. The FTIR method was involved in many studies for structural transition of α-lactalbumin under modified conditions including concentration, pH, heat treatment and pressure (Fang and Dalgleish, 1998; Boye *et al*., 1997; Troullier *et al*., 2000). Thus, the heating induced permanently altered structure by large proportion of modified α-helix components (Fang and Dalgleish, 1998). The presence of the secondary structure components in α-LA as observed by FTIR differs concerning the concentration of the protein in the sample. Thus, at high protein concentration β-sheets are observed as predominant components, however, at low protein concentration predominant secondary structures are unordered structures (Fang and Dalgleish, 1998). The refolding reaction of α-LA after alteration in pH and was successfully observed using stopped flow (SF) FTIR confirming the validity of the method for studying the complex process of proteins' refolding (Troullier *et al*., 2000).

Minor whey proteins

The minor whey proteins were also subjected to structural analysis using the FTIR method. FTIR of Immunoglobulin G (IgG) confirmed great presence of β-sheet and for bovine serum albumin (BSA) it was confirmed high amount of α-helix and β-turn (Bogahawaththa *et al*., 2017; Sathya Devi *et al*., 2011). The observed secondary structures are reduced when these proteins are subjected to shear and heat treatment as observed by intense reorganization of the peak absorbance in the FTIR spectrum. Lactoferrin shown high peak loading for α-helix (1655 cm^{-1}), β-sheets (1636 cm^{-1}) and turns (1677 cm^{-1}) confirming presence of secondary structure components which are greatly affected during temperature and pH treatment (Goulding *et al*., 2021; Geagea *et al*., 2017). Thus, the observed structural transition of the FTIR analysis on whey proteins during changed native conditions aligns with the other findings for crystallographic or NMR analysis confirming that FTIR can be a valuable method for secondary structure observation.

2C.4.2 FTIR studies on caseins

Caseins as most abundant proteins in milk have been common model for secondary structure studies using FTIR analysis. Due to caseins' structural flexibility the Amide I profile was found to have complicated profile as observed with FTIR analysis and in agreement with Raman methodologies (Farrell *et al*., 2013). All four casein (α_{s1}-, α_{s2}-, β-, and κ-CN) and casein micelle have been extensively studied for secondary structure components under native and modified conditions. In addition, smaller fragment and individual polypeptides (casomorphins) that are found isolated from the β-casein molecule have also been studied for structural characteristics. The FTIR studies on each casein, casein micelle and β-casomorphins are discussed individually.

Casein micelle

The secondary structure of casein micelle was resolved using FTIR by Byler and Susi, (1986). The assignment of the structural component in Amide I region of the FTIR spectra were performed by deconvolution and Gaussian peak fitting. The micelles were analysed in deuterium oxide solution and several components were identified including β-sheet (strong) (1621-1639 cm^{-1}), unspecified or random structures (1640-1646 cm^{-1}), helical structures (1651-1657 cm^{-1}), turns (1660-1665 cm^{-1} and 1685-1695 cm^{-1}) and β-structure (weak) (1670-1680 cm^{-1}). The effect of minerals on casein micelle structure was studied by FTIR spectroscopy (Curley *et al*., 1998). The changes in the structural components in aqueous solutions of caseins micelles in presence of calcium (Ca), sodium (Na) and potassium (K) was analysed by Fourier smoothing of the second derivative spectra and curve-fitting of the spectra in Amide I and II region. The study provided global estimation of the structural elements and the hydrophobic forces that held the micellar structure together. Thus, the electrostatic binding of calcium to the caseins leads to redistribution of the FTIR spectra. The presence of K and Na induces decrease in loops and helical structure and increase in turns. The Amide III region of the FTIR spectra of casein micelles in different NaCl concentration was used for structural studies and

correlated to Amide I region (Hussain *et al*., 2011). The spectra from Amide I (1720-1580 cm^{-1}) and Amide III region (1200-1350 cm^{-1}) were baseline corrected, the second derivative was calculated with Satvitsky-Golay smoothing function and spectra were deconvoluted by a non-linear regression curve fitting for Gaussian peaks. The findings from both spectra regions (Amide I and III) identified several components (ß-sheet, irregular, α-helix and turn) that shown high sensitivity on NaCl concentration (Hussain *et al*., 2011). Casein mixtures prepared with different casein ratios in order to imitate micellar conditions were prepared and analysis for structural elements in presence of calcium (Ca) and magnesium (Mg) (Grewal *et al*., 2021). The solvent was subtracted from the spectra and then the spectra were baseline corrected, deconvoluted, smoothed, and second derivative was calculated with Gaussian function and finally PCA analysis was used for better identification of samples groupings and changes in structural components. The study confirmed increase in β-sheets elements and decrease in triple helices and turns when Ca and Mg were added in the solution. The averaged values for the secondary structural components present in individual caseins and casein micelle are presented in table 2.

Table 2. Averaged values of secondary structural components for individual caseins and casein micelle. Updated from Grewal *et al*., 2021; Qi *et al*., 2005; Byler and Susi, (1986).

Average values of structural elements (%)

	α-helix	β-sheets	Random coils	310-helix or PPII	β-turns
α_{s1}- casein	15	38	15	16	13
α_{s2}- casein	14	35	12	15	21
κ-casein	13	33	11	11	22
β-casein	11	30	15	14	18
Casein micelle	14	27	18	NA	41

α_{s1}-Casein

The FTIR analysis of α_{s1}-CN have been carried out on fragments of the molecule. Malin *et al*., (2001) confirmed presence of 30-40 % turns, 25-30 % extended in the N terminal (f 1-23) of the polypeptide chain that was in agreement with CD and NMR data. Additionally, FTIR analysis confirmed presence of polyproline II helix and loops in the α_{s1}-casein molecule. The purified α_{s1}-CN was prepared for FTIR spectroscopy by dilution in 10 mM sodium phosphate buffer and 70 mM NaCl and the obtained spectrum was analysed after applying second derivative and curve-fitting procedure. The C-terminal (f 136-196) was observed to have β-extended (3_{10} structures) (49%), turns (22 %), unordered structures (23 %) and α-helix (5%) (Alaimo *et al*., 1999). The high number of turns in the C-terminal of the molecule is due to the high presence of proline in the polypeptide chain (Malin *et al*., 2001). For this study, the purified α_{s1}-casein was diluted in 25 mM dipotassium PIPES buffer for the FTIR analysis and the spectra was analysed after second derivative, Gaussian fitting and curve fitting. The studies were carried out to observe the possible reactive sites of the molecule.

α_{s2}-Casein

The α_{s2}-CN was studied by FTIR analysis by Hoagland *et al*., (2001). In comparison to the FTIR studies for secondary structure observation performed on α_{s1}-CN where only fragments were observed, here the whole molecule was studied for secondary structure components. The purified α_{s2}-CN was prepared in 25 mM PIPES and 80 mM KCl solution which absorbance in the spectrum was subtracted and the peak resolution was improved by smoothing, second derivative and Gauss-Newton deconvolution (Hoagland *et al*., 2001). From the analysis in the Amide I region was concluded that α_{s2}-CN have 32 % α-helix (including 41-60 residues from the casein chain), 27 % β sheets (extended) and 31% turns (Hoagland *et al*., 2001).

κ-Casein

Kumosinski *et al*., (1991, 1993) studied the structure of κ-CN. In this work, the authors used the spectroscopic data from FTIR results and molecular modelling to predict the secondary structure of κ-CN. The proposed findings were two unstranded β-sheets. Farrell *et al*., (2003) observed the environmental influence on κ-CN structure using FTIR in combination with CD, molecular modelling and electron microscopy. They found that the temperature did not affect the secondary structure gradually and the most observed components were extended strands and β-turns. However, the literature does not provide findings of the secondary structure of κ-CN only obtained by FTIR spectroscopy. Thus, the κ-CN structure was mainly resolved by using other methods including NMR, CD and molecular modelling (Kumosinski *et al*., 1991, 1993; Farrell *et al*., 2003, 1996; Plowman *et al*., 1997; Bansal *et al*., 2006).

β-Casein

The structure of β-CN was extensively studied by FTIR spectroscopy (Farrell Jr *et al*., 2001; Qi *et al*., 2005; Farrell *et al*., 2002; Li *et al*., 2019). Farrell *et al*., (2001) analysed the β-CN for presence of secondary structure components. The study observed structural transition of the molecule at temperature range of 5-70°C. The findings included loss of structural elements at low temperature (cold denaturation), however at increased temperature only the flexible conformational elements (loops, helix and polyproline II) was observed to be influenced by slight increase. The FTIR analysis on β-CN was performed by diluting the molecule in 25 mM dipotassium PIPES buffer and 8 mM KCl. The spectroscopic data was phase-corrected and apodized with Happ-Genzel function and the peaks were identified by non-linear Gaussian regression fitting. The study summarizes several secondary structure components including turns (32%), α-helix/loop (29%), β-sheets (18%) and polyproline II structures (20%). Farrell et al., (2002), continued the study when fragment of β-CN (f1-25) was analysed for changes in structural components concerning changed temperature and phosphorylation. It was confirmed that the polypeptide had stable structural element at temperature level of 5 - 70

°C and little or no changes during phosphorylation. Similarly, as in the previous work by Farrell Jr *et al*., (2001) the samples were prepared in PIPES-KCl buffer and the secondary structure elements were calculated after non-linear regression fitting of Gaussian peaks in Amide I region.

Qi *et al*., (2005) analysed the transition in the secondary structure components resulted by pH and temperature effect using FTIR spectroscopy. The study indicated no changes in the structural components at pH 6.75 and temperature range of 5-70°C. However, major temperature changes were confirmed at high pH (10.5) by increase in the extended structures and decrease in α-helix or loop structures. The presence of polyproline II structures (20%) in β-CN was confirmed in the Amide I region of the FTIR analysis under physiological conditions of the molecule (pH 6.75 and 25 °C) (Qi *et al*., 2005). The estimated secondary structures in physiological conditions obtained by summing the areas of FTIR bands were α-helix/loops (25 %), polyproline II structures (20%), β-sheets (12), turns (29%) and unspecified structures (14%). For this study the authors prepared the purified β-CN in PIPES buffer and the FTIR spectra was phase–corrected with Happ-Genzel function and curve-fitting procedure was applied for better peak assignment. Li *et al*., (2019) observed the structural changes of pure β-CN and β-CN concentrate as a function of temperature and $CaCl_2$ using FTIR spectroscopy. The findings confirmed little change in structural components as a function of temperature that predominately were observed in turns and intramolecular β-sheets elements, and the addition of $CaCl_2$ did not change the structural elements in both pure and concentrated β-CN samples.

Casomorphins

β-casomorphins (BCMs) are opioid peptides that are isolated from β-casein molecule during digestion in the human body an also during milk processing. In the present literature, there are not available studies that use these spectroscopic tools to observe the vibrations of the functional groups of BCMs.

2C.5 Conclusion

The FTIR spectroscopy have been widely implemented on the structural studies on milk proteins. Even though, the FTIR cannot detect the atomic position of the molecules, it can be highly sensitive and advantageous in function and structure related changes in the milk proteins. Thus, the FTIR can successfully determine structural information and the interaction properties of the proteins. The combination of FTIR with chemometric analysis can successfully fingerprint the functional-structural characteristics of milk proteins.

References:

Alaimo, M.H., Farrell Jr, H.M. and Germann, M.W., 1999. Conformational analysis of the hydrophobic peptide αs1-casein (136–196). *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, *1431*(2), pp.410-420.

Bansal, P.S., Grieve, P.A., Marschke, R.J., Daly, N.L., McGhie, E., Craik, D.J. and Alewood, P.F., 2006. Chemical synthesis and structure elucidation of bovine κ-casein (1–44). *Biochemical and biophysical research communications*, *340*(4), pp.1098-1103.

Barth, A., 2007. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1767*(9), pp.1073-1101.

Bewley, M.C., Baker, E.N., Qin, B.Y., Jameson, G.B. and Sawyer, L., 1997. Bovine beta-lactoglobulin and its variants: a three-dimensional structural perspective. In *Milk protein polymorphism, Palmerston North (New Zealand), Feb 1997*. International Dairy Federation.

Bhattacharjee, C., Saha, S., Biswas, A., Kundu, M., Ghosh, L. and Das, K.P., 2005. Structural changes of β-lactoglobulin during thermal unfolding and refolding–an FT-IR and circular dichroism study. *The Protein Journal*, *24*(1), pp.27-35.

Bogahawaththa, D., Chandrapala, J. and Vasiljevic, T., 2017. Thermal denaturation of bovine immunoglobulin G and its association with other whey proteins. *Food Hydrocolloids*, *72*, pp.350-357.

Boye, J.I., Alli, I. and Ismail, A.A., 1997. Use of differential scanning calorimetry and infrared spectroscopy in the study of thermal and structural stability of α-lactalbumin. *Journal of Agricultural and Food Chemistry*, *45*(4), pp.1116-1125.

Boye, J.I., Ismail, A.A. and Alli, I., 1996. Effects of physicochemical factors on the secondary structure of β-lactoglobulin. *Journal of Dairy Research*, *63*(1), pp.97-109.

Braue, E.H. and Pannella, M.G., 1987. Consistency in circle cell FT-IR analysis of aqueous solutions. *Applied spectroscopy*, *41*(6), pp.1057-1067.

Byler, D.M. and Susi, H., 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers: Original Research on Biomolecules*, *25*(3), pp.469-487.

Carbonaro, M. and Nucara, A., 2010. Secondary structure of food proteins by Fourier Transform spectroscopy in the mid-infrared region. *Amino acids*, *38*(3), pp.679-690.

Curley, D.M., Kumosinski, T.F., Unruh, J.J. and Farrell Jr, H.M., 1998. Changes in the secondary structure of bovine casein by Fourier Transform Infrared spectroscopy: Effects of calcium and temperature. *Journal of Dairy Science*, *81*(12), pp.3154-3162.

Daniloski, D., McCarthy, N.A., Markoska, T., Auldist, M.J. and Vasiljevic, T., 2022a. Conformational and physicochemical characteristics of bovine skim milk obtained from cows with different genetic variants of β-casein. *Food Hydrocolloids*, *124*, p.107186.

Daniloski, D., McCarthy, N.A., O'Callaghan, T.F. and Vasiljevic, T., 2022b. Authentication of β-casein milk phenotypes using FTIR spectroscopy. *International Dairy Journal*, *129*, p.105350.

Daniloski, D., McCarthy, N.A. and Vasiljevic, T., 2022c. Impact of heating on the properties of A1/A1, A1/A2, and A2/A2 β-casein milk phenotypes. *Food Hydrocolloids*, *128*, p.107604.

Dong, A., Huang, P. and Caughey, W.S., 1990. Protein secondary structures in water from second-derivative amide I infrared spectra. *Biochemistry*, *29*(13), pp.3303-3308.

Dupont, D., Croguennec, T., Brodkorb, A. and Kouaouci, R., 2013. Quantitation of proteins in milk and milk products. In *Advanced dairy chemistry* (pp. 87-134). Springer, Boston, MA.

Fang, Y. and Dalgleish, D.G., 1998. The conformation of α-lactalbumin as a function of pH, heat treatment and adsorption at hydrophobic surfaces studied by FTIR. *Food Hydrocolloids*, *12*(2), pp.121-126.

Farrell, H.M., Brown, E.M. and Malin, E.L., 2013. Higher order structures of the caseins: A paradox?. In *Advanced dairy chemistry* (pp. 161-184). Springer, Boston, MA.

Farrell, H.M., Cooke, P.H., Wickham, E.D., Piotrowski, E.G. and Hoagland, P.D., 2003. Environmental influences on bovine κ-casein: reduction and conversion to fibrillar (amyloid) structures. *Journal of protein chemistry*, *22*(3), pp.259-273.

Farrell Jr, H.M., Wickham, E.D., Unruh, J.J., Qi, P.X. and Hoagland, P.D., 2001. Secondary structural studies of bovine caseins: temperature dependence of β-casein structure as analyzed by circular dichroism and FTIR spectroscopy and correlation with micellization. *Food Hydrocolloids*, *15*(4-6), pp.341-354.

Farrell, H.M., Qi, P.X., Wickham, E.D. and Unruh, J.J., 2002. Secondary structural studies of bovine caseins: structure and temperature dependence of β-casein phosphopeptide (1-25) as analyzed by circular dichroism, FTIR spectroscopy, and analytical ultracentrifugation. *Journal of Protein Chemistry*, *21*(5), pp.307-321.

Farrell, H.M., Kumosinski, T.F., Cooke, P.H., King, G., Hoagland, P.D., Wickham, E.D., Dower, H.J. and Groves, M.L., 1996. Particle sizes of purified κ-casein: metal effect and

correspondence with predicted three-dimensional molecular models. *Journal of protein chemistry*, *15*(5), pp.435-445.

Gaiani, C., Morand, M., Sanchez, C., Tehrany, E.A., Jacquot, M., Schuck, P., Jeantet, R. and Scher, J., 2010. How surface composition of high milk proteins powders is influenced by spray-drying temperature. *Colloids and Surfaces B: Biointerfaces*, *75*(1), pp.377-384.

Geagea, H., Gomaa, A., Remondetto, G., Moineau, S. and Subirade, M., 2017. Molecular structure of lactoferrin influences the thermal resistance of lactococcal phages. *Journal of Agricultural and Food Chemistry*, *65*(10), pp.2214-2221.

Goulding, D.A., O'Regan, J., Bovetto, L., O'Brien, N.M. and O'Mahony, J.A., 2021. Influence of thermal processing on the physicochemical properties of bovine lactoferrin. *International Dairy Journal*, *119*, p.105001.

Grewal, M.K., Chandrapala, J., Donkor, O., Apostolopoulos, V. and Vasiljevic, T., 2017. Predicting sediment formation in ultra high temperature-treated whole and skim milk using attenuated total reflectance-Fourier Transform infrared spectroscopy. *International Dairy Journal*, *74*, pp.39-48.

Grewal, M.K., Huppertz, T. and Vasiljevic, T., 2018. FTIR fingerprinting of structural changes of milk proteins induced by heat treatment, deamidation and dephosphorylation. *Food Hydrocolloids*, *80*, pp.160-167.

Grewal, M.K., Vasiljevic, T. and Huppertz, T., 2021. Influence of calcium and magnesium on the secondary structure in solutions of individual caseins and binary casein mixtures. *International Dairy Journal*, *112*, p.104879.

Griffiths, P.R., De Haseth, J.A. and Azarraga, L.V., 1983. Capillary Gc/Ft-Ir. *Analytical Chemistry*, *55*(13), pp.1361A-1387A.

Haque, E., Bhandari, B.R., Gidley, M.J., Deeth, H.C. and Whittaker, A.K., 2011. Ageing-induced solubility loss in milk protein concentrate powder: effect of protein conformational modifications and interactions with water. *Journal of the Science of Food and Agriculture*, *91*(14), pp.2576-2581.

Hoagland, P.D., Unruh, J.J., Wickham, E.D. and Farrell Jr, H.M., 2001. Secondary structure of bovine αS2-casein: Theoretical and experimental approaches. *Journal of dairy science*, *84*(9), pp.1944-1949.

Hussain, R., Gaiani, C. and Scher, J., 2011. Revealing casein micelle dispersion under various ranges of NaCl: evolution of particles size and structure. *World Acad Sci Eng Technol*, *73*, pp.972-976.

Iñón, F.A., Garrigues, S. and de la Guardia, M., 2004. Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques. *Analytica Chimica Acta*, *513*(2), pp.401-412.

Karoui, R., Mazerolles, G. and Dufour, É., 2003. Spectroscopic techniques coupled with chemometric tools for structure and texture determinations in dairy products. *International Dairy Journal*, *13*(8), pp.607-620.

Karoui, R., Downey, G. and Blecker, C., 2010. Mid-infrared spectroscopy coupled with chemometrics: a tool for the analysis of intact food systems and the exploration of their molecular structure− quality relationships− a review. *Chemical reviews*, *110*(10), pp.6144-6168.

Kher, A., Udabage, P., McKinnon, I., McNaughton, D. and Augustin, M.A., 2007. FTIR investigation of spray-dried milk protein concentrate powders. *Vibrational Spectroscopy*, *44*(2), pp.375-381.

Kumosinski, T.F. and Farrell Jr, H.M., 1993. Determination of the global secondary structure of proteins by Fourier Transform infrared (FTIR) spectroscopy. *Trends in Food Science & Technology*, *4*(6), pp.169-175.

Kumosinski, T.F., Brown, E.M. and Farrell Jr, H.M., 1991. Three-dimensional molecular modeling of bovine caseins: κ-casein. *Journal of dairy science*, *74*(9), pp.2879-2887.

Kumosinski, T.F., Brown, E.M. and Farrell Jr, H.M., 1993. Three-dimensional molecular modeling of bovine caseins: a refined, energy-minimized κ-casein structure. *Journal of Dairy Science*, *76*(9), pp.2507-2520.

Lefèvre, T. and Subirade, M., 1999. Structural and interaction properties of β-Lactoglobulin as studied by FTIR spectroscopy. *International Journal of Food Science & Technology*, *34*(5-6), pp.419-428.

Li, M., Auty, M.A., Crowley, S.V., Kelly, A.L., O'Mahony, J.A. and Brodkorb, A., 2019. Self-association of bovine β-casein as influenced by calcium chloride, buffer type and temperature. *Food Hydrocolloids*, *88*, pp.190-198.

Malin, E.L., Alaimo, M.H., Brown, E.M., Aramini, J.M., Germann, M.W., Farrell, H.M., McSweeney, P.L. and Fox, P.F., 2001. Solution structures of casein peptides: NMR, FTIR, CD, and molecular modeling studies of αs1-casein, 1–23. *Journal of Protein Chemistry*, *20*(5), pp.391-404.

Markoska, T., Huppertz, T., Grewal, M.K. and Vasiljevic, T., 2019a. FTIR analysis of physiochemical changes in raw skim milk upon concentration. *LWT*, *102*, pp.64-70.

Markoska, T., Huppertz, T., Grewal, M.K. and Vasiljevic, T., 2019b. Structural changes of milk proteins during heating of concentrated skim milk determined using FTIR. *International Dairy Journal*, *89*, pp.21-30.

Mazerolles, G., Devaux, M.F., Duboz, G., Duployer, M.H., Riou, N.M. and Dufour, É., 2001. Infrared and fluorescence spectroscopy for monitoring protein structure and interaction changes during cheese ripening. *Le Lait*, *81*(4), pp.509-527.

Mediwaththe, A., Bogahawaththa, D., Grewal, M.K., Chandrapala, J. and Vasiljevic, T., 2018. Structural changes of native milk proteins subjected to controlled shearing and heating. *Food research international*, *114*, pp.151-158.

Molinari, H., Ragona, L., Varani, L., Musco, G., Consonni, R., Zetta, L. and Monaco, H.L., 1996. Partially folded structure of monomeric bovine β-lactoglobulin. *FEBS letters*, *381*(3), pp.237-243.

Noda, I. and Ozaki, Y., 2005. *Two-dimensional correlation spectroscopy: applications in vibrational and optical spectroscopy*. John Wiley & Sons.

Pelton, J.T. and McLean, L.R., 2000. Spectroscopic methods for analysis of protein secondary structure. *Analytical biochemistry*, *277*(2), pp.167-176.

Plowman, J.E., Creamer, L.K., Liddell, M.J. and Cross, J.J., 1997. Solution conformation of a peptide corresponding to bovine κ-casein B residues 130–153 by circular dichroism spectroscopy and 1H-nuclear magnetic resonance spectroscopy. *Journal of dairy research*, *64*(3), pp.377-397.

Prestrelski, S.J., Byler, D.M. and Thompson, M.P., 1991. Infrared spectroscopic discrimination between α-and 310-helices in globular proteins: Reexamination of Amide I infrared bands of α-lactalbumin and their assignment to secondary structures. *International Journal of Peptide and Protein Research*, *37*(6), pp.508-512.

Qi, X.L., Holt, C., Mcnulty, D., Clarke, D.T., Brownlow, S. and Jones, G.R., 1997. Effect of temperature on the secondary structure of β-lactoglobulin at pH 6.7, as determined by CD and IR spectroscopy: a test of the molten globule hypothesis. *Biochemical Journal*, *324*(1), pp.341-346.

Qi, P.X., Wickham, E.D., Piotrowski, E.G., Fagerquist, C.K. and Farrell, H.M., 2005. Implication of C-terminal deletion on the structure and stability of bovine β-casein. *The protein journal*, *24*(7), pp.431-444.

Qin, B.Y., Bewley, M.C., Creamer, L.K., Baker, H.M., Baker, E.N. and Jameson, G.B., 1998. Structural basis of the Tanford transition of bovine β-lactoglobulin. *Biochemistry*, *37*(40), pp.14014-14023.

Rahaman, T., Vasiljevic, T. and Ramchandran, L., 2015. Conformational changes of β-lactoglobulin induced by shear, heat, and pH—Effects on antigenicity. *Journal of dairy science*, *98*(7), pp.4255-4265.

Rodriguez-Saona, L., Ayvaz, H. and Wehling, R.L., 2017. Infrared and Raman spectroscopy. In *Food analysis* (pp. 107-127). Springer, Cham.

Sathya Devi, V., Coleman, D.R. and Truntzer, J., 2011. Thermal unfolding curves of high concentration bovine IgG measured by FTIR spectroscopy. *The Protein Journal*, *30*(6), pp.395-403.

Sjaunja, L.O. and Andersson, I., 1985. Laboratory experiments with a new infrared (IR) milk analyzer, the Milko-Scan 605. *Acta agriculturae scandinavica*, *35*(4), pp.345-352.

Susi, H. and Byler, D.M., 1983. Protein structure by Fourier Transform infrared spectroscopy: second derivative spectra. *Biochemical and biophysical research communications*, *115*(1), pp.391-397.

Troullier, A., Reinstädler, D., Dupont, Y., Naumann, D. and Forge, V., 2000. Transient non-native secondary structures during the refolding of α-lactalbumin detected by infrared spectroscopy. *Nature structural biology*, *7*(1), pp.78-86.

Uhrínová, S., Uhrín, D., Denton, H., Smith, M., Sawyer, L. and Barlow, P.N., 1998. Complete assignment of 1H, 13C and 15N chemical shifts for bovine β-lactoglobulin: secondary structure and topology of the native state is retained in a partially unfolded form. *Journal of Biomolecular NMR*, *12*(1), pp.89-107.

Van de Voort, F.R., 1992. Fourier Transform infrared spectroscopy applied to food analysis. *Food Research International*, *25*(5), pp.397-403.

Yang, H., Yang, S., Kong, J., Dong, A. and Yu, S., 2015. Obtaining information about protein secondary structures in aqueous solution using Fourier Transform IR spectroscopy. *Nature protocols*, *10*(3), pp.382-396.

Ye, M.P., Zhou, R., Shi, Y.R., Chen, H.C. and Du, Y., 2017. Effects of heating on the secondary structure of proteins in milk powders using mid-infrared spectroscopy. *Journal of Dairy Science*, *100*(1), pp.89-95.

Zhou, Q., Sun, S.Q., Yu, L., Xu, C.H., Noda, I. and Zhang, X.R., 2006. Sequential changes of main components in different kinds of milk powders using two-dimensional infrared correlation analysis. *Journal of Molecular Structure*, *799*(1-3), pp.77-84.

CHAPTER 3: pH-induced changes in β-casomorphin 7 structure studied by 1H Nuclear Magnetic Resonance and Fourier-Transform infrared spectroscopy

DocuSign Envelope ID: C7AFBA28-8EE3-42F6-BADC-B4D1B5ADAD78

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book: pH-induced changes in β-casomorphin 7 structure studied by 1H nuclear magnetic resonance and Fourier-transform infrared spectroscopy.

Surname: Markoska First name: Tatijana

Institute: Institute for Sustainable Industries and Liveat Candidate's Contribution (%): 65

Status:

Accepted and in press: ✔ Date: 19/05/2021

Published: ✔ Date: 28/05/2021

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Tatijana Markoska Digitally signed by Tatijana Markoska Date: 2023.03.05 18:50:48 +11'00' | 05/03/2023

Signature **Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne),
02475D (Sydney), RTO 3113

DocuSign Envelope ID: C7AFBA28-8EE3-42F6-BADC-B4D1B5ADAD78

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

College of Health and Biomedicine, Victoria University, Werribee Campus, Melbourne, Victoria, Australia.

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Tatijana Markoska	85	Experiment design and execution,data analysis, manuscript drafting, revision	[redacted]	30/3/23
Thom Huppertz	5	Designing experiments, manuscript editing and revision		3/30/2023
Todor Vasiljevic	10	Designing experiments, manuscript editing and revision and journal submission		30/03/2023

Updated: September 2019

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83 776954731
CRICOS Provider No. 00124K (Melbourne)
02475D (Sydney), RTO 3113

International Dairy Journal 121 (2021) 105106

Contents lists available at ScienceDirect

International Dairy Journal

journal homepage: www.elsevier.com/locate/idairyj

pH-induced changes in β-casomorphin 7 structure studied by ^{1}H nuclear magnetic resonance and Fourier-transform infrared spectroscopy

Tatijana Markoska [a], Thom Huppertz [a, b, c], Todor Vasiljevic [a, *]

[a] Advanced Food Systems Research Unit, Institute for Sustainable Industries & Liveable Cities and College of Health and Biomedicine, Victoria University, Melbourne, VIC 8001, Australia
[b] FrieslandCampina, Amersfoort, The Netherlands
[c] Wageningen University & Research, Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 3 February 2021
Received in revised form
16 May 2021
Accepted 19 May 2021
Available online 28 May 2021

ABSTRACT

β-Casomorphins are opioid peptides that can be liberated from β-casein during digestion, with some suggested effects on human physiology linked to their properties. In this study, the structural behaviour of β-casomorphin 7 (BCM7; Tyr–Pro–Phe–Pro–Gly–Pro–Ile) was studied at pH 2.3 and 6.7 using ^{1}H nuclear magnetic resonance (^{1}H NMR) and Fourier-transform infrared (FTIR) spectroscopy. *Cis–trans* isomerism involving the Tyr^{1}–Pro^{2}, Phe^{3}–Pro^{4} and Gly^{5}–Pro^{6} bonds of BCM7 was observed. *Cis* and *trans* isomers of Tyr^{1}–Pro^{2} and Phe^{3}–Pro^{4} and *trans* isomers for Gly^{5}–Pro^{6} were observed at both pH 6.7 and 2.3 but *cis* isomers of Gly^{5}–Pro^{6} were only observed at pH 2.3 and not at pH 6.7. These pH-induced conformational changes at the Gly^{5}–Pro^{6} bond of BCM7 may play important role in enzymatic release of this peptide from the parent protein β-casein under gastro-intestinal conditions, as well as further cleavage of BCM7 into smaller peptides.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

β-Casomorphins (BCMs) are a group of opioid peptides that can be released by enzymatic digestion of β-casein. They share the Tyr–X–Phe sequence at the N-terminus with other μ-receptor agonists. A number of BCMs can be released during digestion of β-casein, varying in length between 4 and 11 amino acid residues. The BCM peptide containing 7 amino acid residues, frequently termed BCM7, has been one of the most examined BCMs due to suggested relevance in physiological pathways (Thiruvengadam, Venkidasamy, Thirupathi, Chung, & Subramanian, 2021). This peptide, which has the sequence Tyr–Pro–Phe–Pro–Gly–Pro–Ile, relates to positions 60–66 in the parent protein β-casein and is present in all known genetic variants of bovine β-casein. Genetic variation, however, does occur at position 67 in the bovine β-casein molecule; β-casein A1, B and F contain His, whereas β-casein A2, A3 and I contain Pro at position 67 (Huppertz, 2013). These subclasses of β-casein are commonly referred to as the A1 family and A2 family, respectively. The His versus Pro difference at position 67 has been considered to strongly affect the hydrolysis of the peptide bond between positions 66 and 67 in β-casein (Nguyen, Busetti, Smolenski, Johnson, & Solah, 2021). While some studies suggested that the Ile^{66}–Pro^{67} bond in β-caseins of the A2 family was resistant to hydrolysis under (simulated) gastro-intestinal conditions, thus preventing the release of BCM7 from β-casein A2 (De Noni, 2008; Jinsmaa & Yoshikawa, 1999), several recent studies have shown that the peptide is also released from β-casein A2, A3 and I, and only some quantitative difference in release is observed between proteins from the β-casein A1 and A2 family (Asledottir et al., 2017; Cieślińska, Kamiński, Kostyra, & Sienkiewicz-Szłapka, 2007; Lambers, Broeren, Heck, Bragt, & Huppertz, 2021; Nguyen et al., 2021). Release of BCM7 has also been suggested to be affected by the product matrix and the processing that it has undergone (Lambers et al., 2021; Nguyen, Johnson, Busetti, & Solah, 2015).

While all peptide bonds can show *cis–trans* isomerism, most peptide bonds typically and almost exclusively (>99.9%) adopt the *trans* isomer. This is related to the fact that the amide hydrogen in the *trans* isomer offers less steric repulsion for the preceding C_{α} atom than does the following C_{α} atom. Pro-containing peptide-

* Corresponding author. Tel.: +61 3 9919 8062.
E-mail address: todor.vasiljevic@vu.edu.au (T. Vasiljevic).

https://doi.org/10.1016/j.idairyj.2021.105106
0958-6946/© 2021 Elsevier Ltd. All rights reserved.

bonds, however, are the exception, with notable levels of *cis* isomers also readily observable due to a comparatively small free energy difference between both isomers (Wedemeyer, Welker, & Scheraga, 2002). The presence of 3 Pro residues in BCM7 creates the possibility for up to 8 conformational isomers, which can impact peptide interactions and further enzymatic degradation of the peptide. Previous work by Basosi, D'Amelio, Gaggelli, Pogni, and Valensin (2001) indicated that not all isomers of BCM7 were binding copper and that conformational states of the X—Pro bonds were shown to strongly affect the copper binding by adjacent residues. Despite the acclaimed physiological importance of BCM7, it is surprising that little further attention has been paid to BCM7 isomerisation and its impact on potential physiological activity. In this study, we investigated the effect of pH on BCM7 structure. For this purpose, we selected pH 2.3 and 6.7, which provide insights into the behaviour of the peptide at gastric and intestinal conditions. ^{1}H nuclear magnetic resonance (NMR) and Fourier-transform infrared (FTIR) spectroscopy were used as powerful techniques for this peptide structural studies.

2. Materials and methods

2.1. Sample preparation

The peptide of BCM7 was synthesised by Mimotopes (Mulgrave, Australia) and was of analytical reagent grade (purity ≥ 98%) and used without further purification. The peptide was prepared for analysis by dispersing 10 mg of the peptide in 1 mL H_2O:D_2O (90:10%). The pH was adjusted with to pH 2.3 with 0.25 M HCl or to pH 6.7 with 0.25 M NaOH.

2.2. Fourier transform infrared spectroscopy

The secondary structure of BCM7 at pH 2.3 and 6.7 was analysed using a PerkinElmer FTIR spectrometer (Frontier, PerkinElmer, Boston, MA, USA) in the range of 4000 to 600 cm^{-1} with a resolution of 4 cm^{-1} and averaging 16 scans for each spectrum. At the start of measurement, the background spectrum was scanned with a blank diamond attenuated total reflectance (ATR) cell. Before the analysis the solvent spectrum (90%H_2O/10%D_2O) was recorded and used for subtraction from the sample spectra to eliminate the intense solvent signal. The original solvent-subtracted spectrum in the region 1800–1100 cm^{-1} was analysed for C=O stretching (amide I) and C–N stretching and N–H bending (amide II and III). The spectra analysis was performed using OriginPro 2020b software (OriginLab, Northampton, MA, USA). For the amide I region, Savitzky–Golay smoothing was used. The baseline of averaged spectra from five repetitions was subtracted and curve fitting was applied on amide I region using Gaussian distribution with R^2 of 0.998 and 0.997 for the sample with pH 2.3 and 6.7, respectively.

2.3. Nuclear magnetic resonance

The ^{1}H NMR experiments were performed on a Bruker Avance spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany) operating at a 600 MHz transmitter frequency using 5 mm TXI probe with Z- and XYZ-gradient. The NMR spectra were acquired at a temperature of 295 K. The accumulated proton NMR was acquired using 16 scans and spectra width of 9615 Hz. The two-dimensional spectra that were accumulated included phase sensitive double quantum-filtered correlation spectroscopy (DQF-COSY), total correlation spectroscopy (TOCSY) and rotating frame Overhause effect spectroscopy (ROESY). The DQF-COSY spectra were recorded using 8 scans and a spectral width of 8196 Hz. The TOCSY spectra used Homonuclear-Hahn transfer spectroscopy and DIPSI2 sequence for mixing (Bax & Davis, 1985) and was recorded using 8 scans and a spectral width of 8196 Hz. The ROESY spectrum with continuous wave spinlock for mixing was acquired using 8 scans and a spectral width of 5882 Hz. In all spectra, water suppression was achieved using excitation sculpting with gradients (Hwang & Shaka, 1995). The acquisition mode used for two-dimensional spectra was States-TPPI (time-proportional phase incrementation) (Marion & Wüthrich, 1983). NMR data processing was performed using TopSpin (version 4.0.6) software (Bruker BioSpin). The Field intensity decays were corrected with a 0.3 Hz line-broadening parameter. All spectra were phase-corrected using 0th and 1st order correction for pk. Chemical shift assignment was accomplished using corresponding amino acid atom values obtained by Wüthrich (1986) and from BMRB files (Seavey, Farr, Westler, & Markley, 1991).

3. Results and discussion

In the NMR spectra, the chemical shift (δ) assignment was performed using TOCSY and ROESY NMR. TOCSY NMR spectra provides information of the position of ^{1}H from neighbouring amino acids connected through bound (Markoska, Vasiljevic, & Huppertz, 2020). However, for structural elucidation, ROESY spectra were used with the aim of observing the through space coupling of the protons. In the TOCSY spectra, the alpha-amide region ($J_{\alpha N}$) or fingerprint region of amino acids is presented (Fig. 1A,B). The position of the $J_{\alpha N}$ of each amino acid was analysed using previous assignment strategies (Wüthrich, 1986). The TOCSY spectra were used to detect the chemical shifts of Tyr^{1}, Phe^{3}, Gly^{5} and Ile^{7} in the fingerprint or amide (HN) region (Fig. 1A,B). Compared with pH 2.3, the peptide at pH 6.7 promoted upfield shifting of the HN chemical shifts of Ile, which shielded for 0.6 ppm. The proton shielding can be observed also in the overlapped ^{1}H NMR spectra of BCM7 at pH 2.3 and 6.7 (Fig. 1C). The large NH peak appeared as a triplet at pH 2.3, but at pH 6.7 appeared as a doublet shielding upfield. The ^{1}H NMR shielding is a result of changes in proton distribution as a function of the changed environment (Shiga, Suzuki, & Tachikawa, 2010). The upfield shielding of NH indicates the presence of *cis* conformers in the system (Alderson, Lee, Charlier, Ying, & Bax, 2018). In addition, in the original FTIR spectra, a broad peak at 1519 cm^{-1} at pH 2.3 was observed that moved to positive absorbance and separated into two peaks at pH 6.7 (Fig. 2A). This peak previously was assigned to a NH-symmetric bending and symmetric rock of Ile (Hernandez, Pflüger, Nsangou, & Ghomi, 2009). Therefore, the vibrational bending of NH observed in the given region of the FTIR spectra can be related to NH bending vibration of the amino acids. Moreover, the NH group of amino acids in BCM7 was affected by an increase in pH, which promoted intense proton exchanges and structural vibrations. This confirms proton exchange and structural vibration of NH from Ile^{7} as observed in the NMR spectra.

The $-CH_2-$ $(CH_3)_2$ side chain of Ile^{7} also slightly shifted up field as a result of pH adjustment from acid to neutral (Fig. 1C). For the other amino acids, only slight shielding was observed in the amide and alpha protons; however, the sidechains of Tyr^{1} and Phe^{3} were not affected by differences in pH. Hence, the aromatic region in the NMR spectra (6.5–7.5 ppm) shown nearly identical peak position and intensity. The NMR chemical shifts are pH dependent as a result of protonation/deprotonation behaviour and structural vibration of the peptides (Christl & Roberts, 1972). Thus, the observed shielding of NH and side chains of Ile at pH 6.7 is a result of vibrational bending in this part of the molecule. In addition, in the FTIR spectra the region 1385–1395 cm^{-1} moved to negative value as pH was increased from 2.3 to 6.7 (Fig. 2A). This change likely resulted from bending vibrations of $-CH_2(CH_3)_2$ (Hernandez, Pflüger, Nsangou, & Ghomi, 2009; Moorthi, Gunasekaran, &

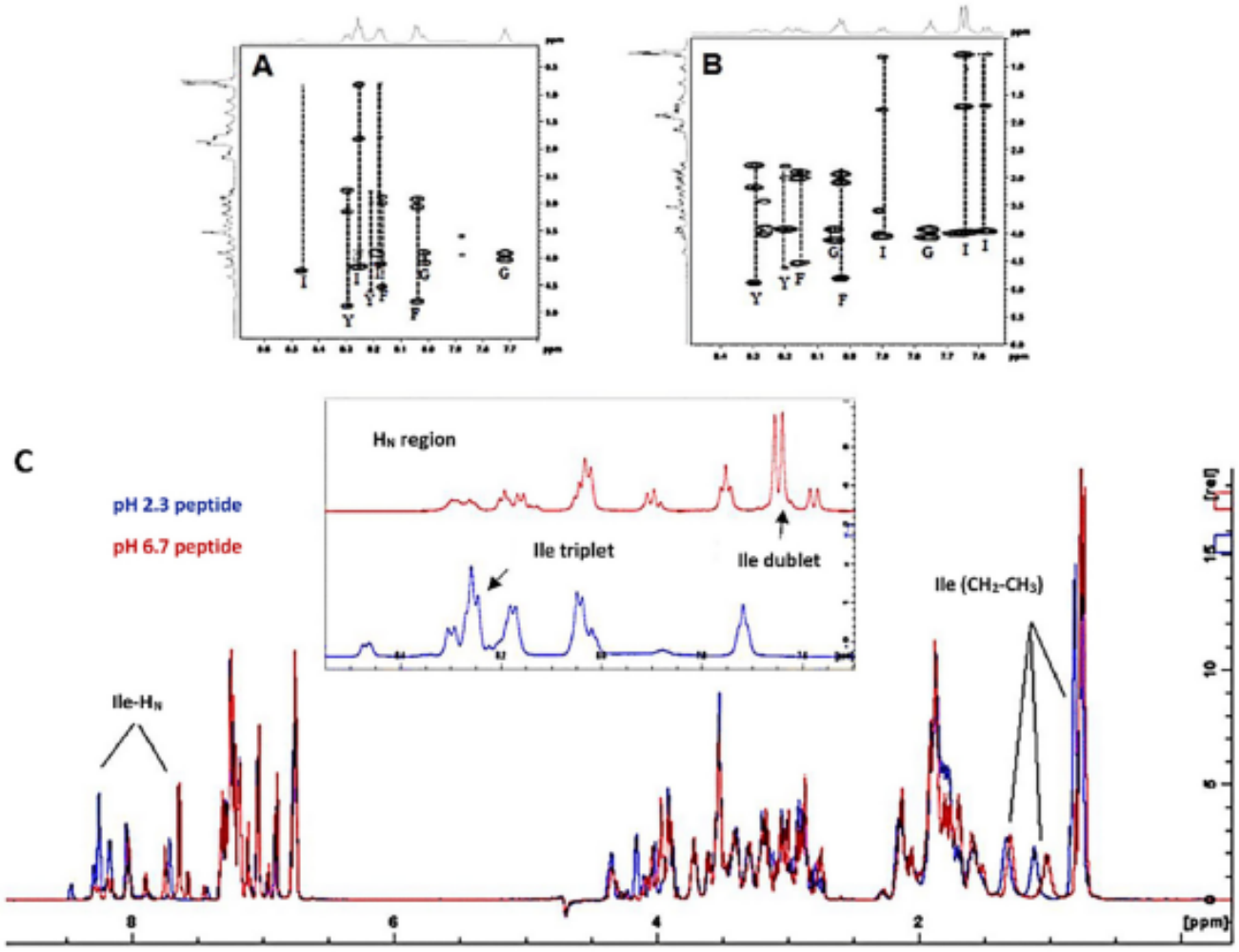


Fig. 1. TOCSY region of the NMR spectra of peptide at pH 2.3 (A) and peptide at pH 6.7 (B); overlapped ^{1}H NMR of peptides with pH 2.3 and 6.7 (C).

Ramkumaar, 2014). The carboxyl group of the Ile in the peptide was not detected in the NMR spectra at either pH due to intense proton exchange with deuterium. The chemical shifts of all observed protons of samples with pH 2.3 and 6.7 are shown in Supplementary material Table S1.

Proline, due to lack of amide group, cannot be assigned using TOCSY NMR and its chemical shift was observed using DQF-COSY and ROESY spectra (Supplementary material Table S1). In the ROESY spectra, the alpha-amide region ($J_{\alpha N}$) was used to assign the dipolar coupling of alpha protons of Pro residues and amide protons in the amino acids (Fig. 3). At both pH values, the $J_{N\alpha}$ of the most intense peaks for Phe^1, Gly^3 and Ile^7 were in close proximity, i.e., within 5 Å, of α protons of Pro-residues. In addition to the most intense peaks, low intensity peaks and their corresponding cross peaks were observed for each amino acid in the peptide. These additional peaks confirmed presence of isomerism of the peptide bonds and thus different conformations of the peptide. For both pH 2.3 and 6.7, a sequential assignment of BCM7 was achieved using sequential NOEs; thus, *cis* and *trans* isomers were distinguished. Previously the existence of eight possible *cis* and *trans* isomerism of BCM7 was confirmed (Basosi et al., 2001). In the current work we observed only four possible isomers for BCM7 at pH 2.3, whereas eight isomers were observed at pH 6.7 (Table 1). This is direct evidence that the Gly^5–Pro^6 bond of BCM7 is predominately in *trans* position at pH 2.3, but on increasing pH to 6.7 also isomerises to *cis* conformers. The FTIR region 1400–1450 cm^{-1} is related to ring vibrations of Pro residues (Barth, 2000). Presented in Fig. 2A, this region shows a strong positive peak at pH 2.3, which was lost when pH was adjusted to 6.7. This spectral region shows high loading due to presence of three Pro residues in the peptide chain. In the sample at pH 6.7, the existence of new isomers promoted bending of the peptide and reorientation of Pro rings. Thus, FTIR confirmed the structural variation in Pro residues as observed in NMR spectra. In addition, from the curve fitting of amide I region (Fig. 2B) the observed high peak at 1678 cm^{-1} for the peptide at both pH values is indicative of *trans* stretching (Vien, Colthup, Fateley, &). This peak shows greater loading at pH 2.3 than pH 6.7, suggesting a greater presence of *trans* isomers at pH 2.3. The peak at 1620 cm^{-1} was more prominent in the sample at pH 6.7 indicating greater presence of *cis* structures (Vien et al., 1991). This *cis–trans* isomerism of peptide-bonds has previously been indicated to be important for the hydrolytic preferences of, e.g., trypsin and pepsin (Lin & Brandts, 1985; Vance, LeBlanc, & London, 1997). Both enzymes hydrolyse X–Pro bonds more intensively when it exists in the *trans* position. The possible combination of isomers identified by NOEs occurrence is presented in Table 1.

The presence of Pro residues in the protein chain impedes the defined structure in caseins due to *cis–trans* isomerism. In BCM7,

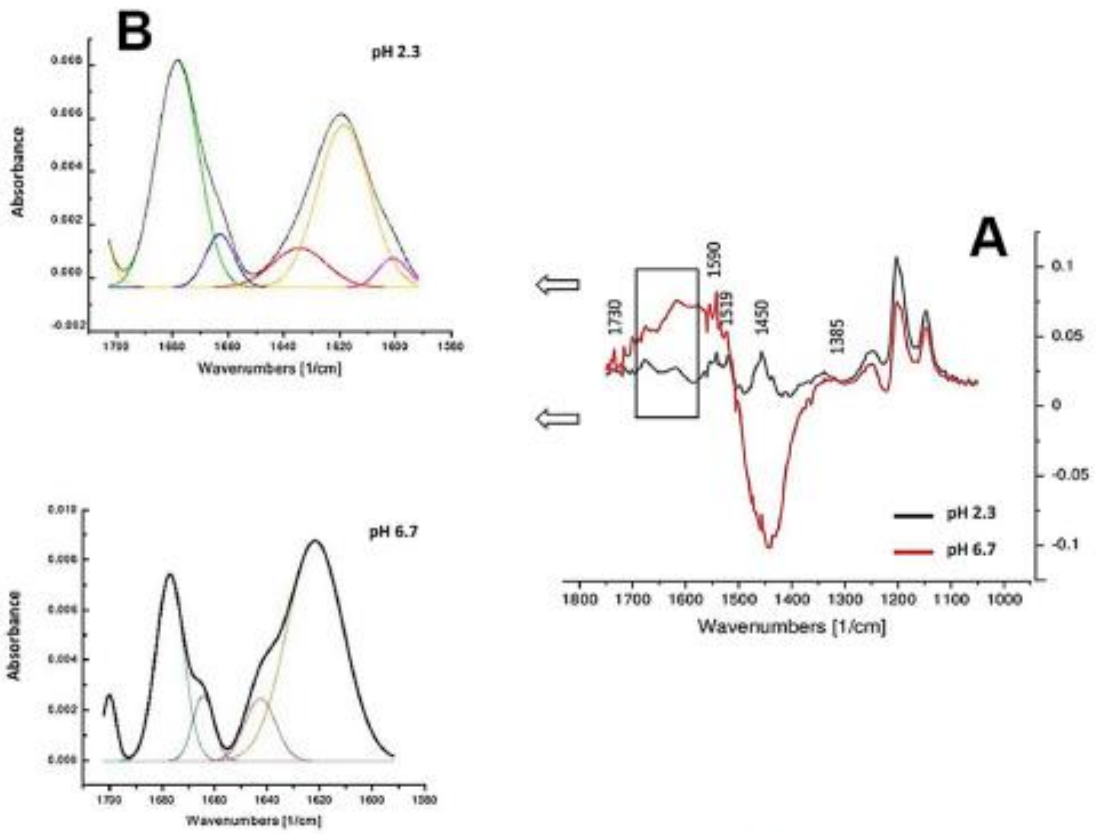


Fig. 2. The original solvent-subtracted FTIR spectra in the region from 1000 to 1750 cm^{-1} (amide I and II) for peptide with pH 2.3 (black) and pH 6.7 (red) (A); curve fitting of FTIR spectra from amide I region for both peptides with cumulative peak (black) and individual peaks (pink, orange, red, blue and green) (B). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

three prolines are present, which was confirmed in this work in addition to confirmed existence of different isomerism, which resulted from a change in the pH environment. The Gly^5–Pro^6–Ile^7 region of the peptide is more effected by pH due to isomerism at the Gly^5–Pro^6 peptide bond. The presence of Pro residues in the peptide structure was previously shown to be active β-turn inducer

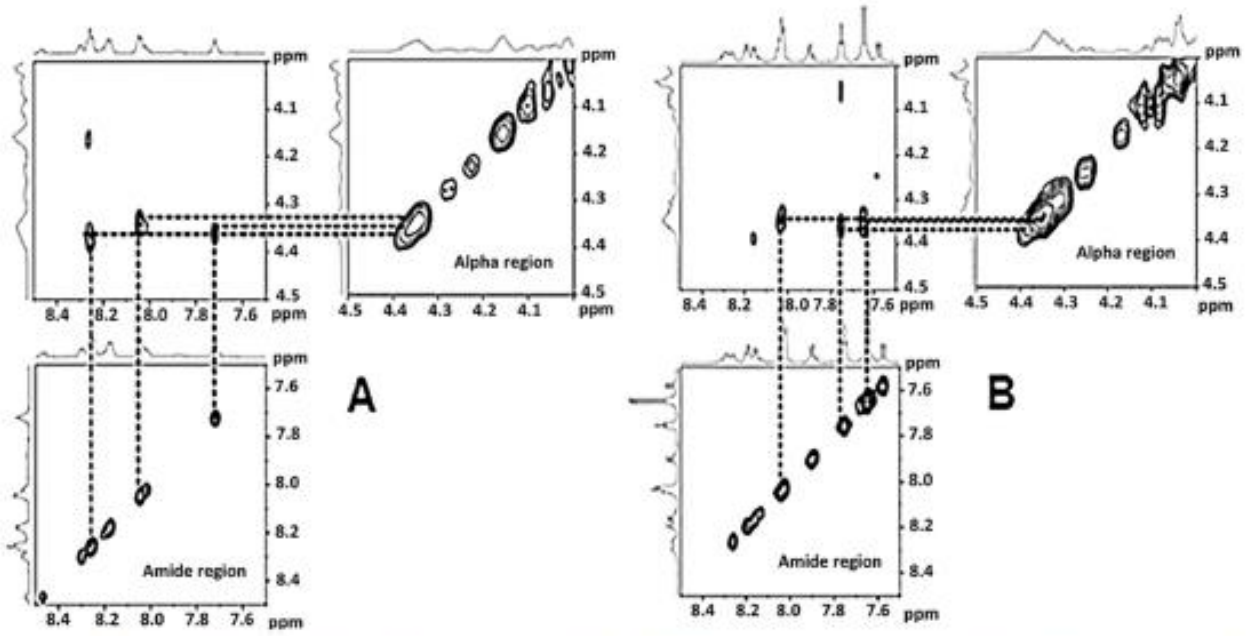
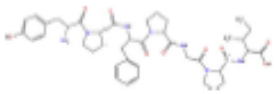
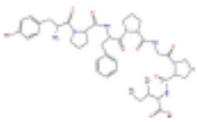
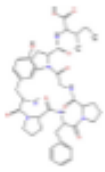
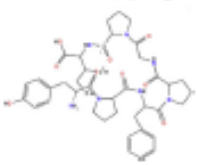
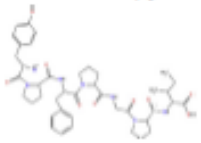
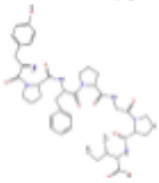
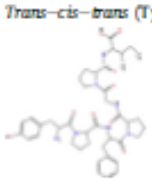
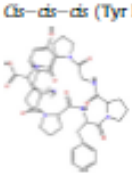


Fig. 3. NOE cross peaks of alpha protons (H_α) of proline and amide protons (H_N) of phenylalanine, glycine and isoleucine in peptide at pH 2.3 (A) and pH 6.7 (B).

Table 1
Trans–cis conformers of β casomorphin at pH 2.3 and 6.7.

Isomer structure	pH	Isomer structure	pH
Trans–trans–trans (Tyr H_{α}-^{1}Pro $H_{\delta\delta}$ > Phe H_{α}-^{2}Pro $H_{\beta\beta}$ > Gly H_{α}-^{3}Pro$H_{\delta\delta}$)	2.3, 6.7	*Trans–trans–cis* (Tyr H_{α}-^{1}Pro $H_{\delta\delta}$ > Phe H_{α}-^{2}Pro $H_{\delta\delta}$ > Gly H_{α}-^{3}Pro$_{\alpha}$)	6.7
			
Cis–cis–trans (Tyr H_{α}-Pro H_{α} > Phe H_{N}-^{2}Pro H_{α} > Gly H_{α}-^{3}Pro$H_{\delta\delta}$)	2.3, 6.7	*Trans–cis–cis* (Tyr H_{α}-^{1}Pro $H_{\delta\delta}$ > Phe H_{N}-^{2}Pro H_{α} > Gly H_{α}-^{3}Pro$_{\alpha}$)	6.7
			
Cis–trans–trans (Tyr H_{α}-Pro H_{α} > Phe H_{α}-^{2}Pro $H_{\delta\delta}$ > Gly H_{α}-^{3}Pro$H_{\delta\delta}$)	2.3, 6.7	*Cis–trans–cis* (Tyr H_{α}-Pro H_{α} > Phe H_{α}-^{2}Pro $H_{\delta\delta}$ > Gly H_{α}-^{3}Pro$_{\alpha}$)	6.7
			
Trans–cis–trans (Tyr H_{α}-^{1}Pro $H_{\delta\delta}$ > Phe H_{N}-^{2}Pro H_{α} > Gly H_{α}-^{3}Pro$H_{\beta\beta}$)	2.3, 6.7	*Cis–cis–cis* (Tyr H_{α}-Pro H_{α} > Phe H_{N}-^{2}Pro H_{α} > Gly H_{α}-^{3}Pro$_{\alpha}$)	6.7
			

(Farahani et al., 2014). Moreover, the side chains of Ile^{7} promoted upfield shifting which is a result of reorientation of this part of the molecule due to chain bending induced by presence of *cis* isomers of X–Pro^{6} bond (Farahani et al., 2014). This part of BCM7 may have an important role in structural behaviour during enzymatic cleavage and biological activity of the peptide. BCM7 has been shown to be an opioid agonist that binds to the μ-receptors in the human body and leads to a range of physiological responses (Daniloski et al., 2021). In addition, the peptide further cleavage and its inactivation is mediated by dipeptidyl peptidase-4 (DPP-4) (Jarmołowska et al., 2007). DPP-4 was shown to have an optimal pH activity at 7.9 (Tereshchenkova et al., 2016), thus peptide isomerism as a result of pH change may play an important role in its activation and deactivation in the human body, which should be further investigated.

4. Conclusion

This study confirmed that combination of NMR, FTIR and chemometrics can be powerful combination for structural evaluation of the peptides. Multidimensional NMR provided detailed information of the proton positions and FTIR supported the predominant structural features in the peptide. The *cis–trans* isomerism of BCM7 as a result of different pH environments was studied presenting predominant *trans* conformers and increase of *cis* isomerism at pH 6.7. These subtle changes appear at the Gly^{5}–Pro^{6} bond that has an important role due to enzymatic hydrolysis of the peptide at this part of the molecule. In addition, cleavage of β-casein into BCM peptides (BCM15, BCM11, BCM7, BCM5) is induced by proteolytic enzymes (pepsin, trypsin, elastase, etc) in the human gut and intestinal lumen (Daniloski et al., 2021). The activity of enzymes to cleave polypeptides into smaller fragments depends on the pH environment and accessibility of the free peptide bonds. The *cis–trans* isomerism of X–Pro bond may be significant indication for BCM activity in human body. Therefore, further studies should follow on structural variations of larger peptides.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The financial and technical support for this study was provided by Victoria University in form of an International Postgraduate Program Scholarship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.idairyj.2021.105106.

References

Alderson, T. R., Lee, J. H., Charlier, C., Ying, J., & Bax, A. (2018). Propensity for cis-proline formation in unfolded proteins. *ChemBioChem*, 19, Article 37.
Asledottir, T., Le, T. T., Petrat-Melin, B., Devold, T. G., Larsen, L. B., & Vegarud, G. E. (2017). Identification of bioactive peptides and quantification of β-casomorphin-7 from bovine β-casein A1, A2 and I after ex vivo gastrointestinal digestion. *International Dairy Journal*, 71, 98–106.
Barth, A. (2000). The infrared absorption of amino acid side chains. *Progress in Biophysics and Molecular Biology*, 74, 141–173.

Basosi, R., D'Amelio, N., Gaggelli, E., Pogni, R., & Valensin, G. (2001). cis–trans Isomerization of β-casomorphin peptides bound to copper (II): integration of EPR and NMR studies. *Journal of the Chemical Society, Perkin Transactions, 2*, 252–257.

Bax, A., & Davis, D. G. (1985). MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *Journal of Magnetic Resonance, 65*, 355–360.

Christl, M., & Roberts, J. D. (1972). Nuclear magnetic resonance spectroscopy. Carbon-13 chemical shifts of small peptides as a function of pH. *Journal of the American Chemical Society, 94*, 4565–4573.

Cieślińska, A., Kamiński, S., Kostyra, E., & Sienkiewicz-Szłapka, E. (2007). Beta-casomorphin 7 in raw and hydrolyzed milk derived from cows of alternative β-casein genotypes. *Milchwissenschaft, 62*, 125–127.

Daniloski, D., Cunha, N. M., McCarthy, N. A., O'Callaghan, T. F., McParland, S., & Vasiljevic, T. (2021). Health-related outcomes of genetic polymorphism of bovine β-casein variants: A systematic review of randomised controlled trials. *Trends in Food Science & Technology, 111*, 233–248.

De Noni, I. (2008). Release of β-casomorphins 5 and 7 during simulated gastrointestinal digestion of bovine β-casein variants and milk-based infant formulas. *Food Chemistry, 110*, 897–903.

Farahani, M. D., Honarparvar, B., Albericio, F., Maguire, G. E., Govender, T., Arvidsson, P. I., et al. (2014). Proline N-oxides: Modulators of the 3D conformation of linear peptides through "NO-turns". *Organic and Biomolecular Chemistry, 12*, 4479–4490.

Hernandez, B., Pflüger, F., Nsangou, M., & Ghomi, M. (2009). Vibrational analysis of amino acids and short peptides in hydrated media. IV. Amino acids with hydrophobic side chains: l-Alanine, l-valine, and l-isoleucine. *Journal of Physical Chemistry B, 113*, 3169–3178.

Huppertz, T. (2013). Chemistry of the caseins. In P. L. H. McSweeney, & P. F. Fox (Eds.) (4th ed.*Proteins: Basic aspects: Vol. 1A. Advanced dairy chemistry* (pp. 135–160). Boston, MA, USA: Springer.

Hwang, T. L., & Shaka, A. J. (1995). Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *Journal of Magnetic Resonance, Series A, 112*, 275–279.

Jarmołowska, B., Bielikowicz, K., Iwan, M., Sidor, K., Kostyra, E., & Kaczmarski, M. (2007). Serum activity of dipeptidyl peptidase IV (DPPIV; EC 3.4.14.5) in breast-fed infants with symptoms of allergy. *Peptides, 28*, 678–682.

Jinsmaa, Y., & Yoshikawa, M. (1999). Enzymatic release of neocasomorphin and β-casomorphin from bovine β-casein. *Peptides, 20*, 957–962.

Lambers, T. T., Broeren, S., Heck, J., Bragt, M., & Huppertz, T. (2021). Processing affects beta-casomorphin peptide formation during simulated gastrointestinal digestion in both A1 and A2 milk. *International Dairy Journal, 121*, 105099.

Lin, L. N., & Brandts, J. F. (1985). Isomer-specific proteolysis of model substrates: Influence that the location of the proline residue exerts on cis trans specificity. *Biochemistry, 24*, 6533–6538.

Marion, D., & Wüthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H–^1H spin-spin coupling constants in proteins. *Biochemical and Biophysical Research Communications, 113*, 967–974.

Markoska, T., Vasiljevic, T., & Huppertz, T. (2020). Unravelling conformational aspects of milk protein structure - Contributions from nuclear magnetic resonance studies. *Foods, 9*, Article 1128.

Moorthi, P. P., Gunasekaran, S., & Ramkumaar, G. R. (2014). Vibrational spectroscopic studies of isoleucine by quantum chemical calculations. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 124*, 365–374.

Nguyen, D. D., Busetti, F., Smolenski, G., Johnson, S. K., & Solah, V. A. (2021). Release of beta-casomorphins during in-vitro gastrointestinal digestion of reconstituted milk after heat treatment. *LWT, 136*, Article 110312.

Nguyen, D. D., Johnson, S. K., Busetti, F., & Solah, V. A. (2015). Formation and degradation of beta-casomorphins in dairy processing. *Critical Reviews in Food Science and Nutrition, 55*, 1955–1967.

Seavey, B. R., Farr, E. A., Westler, W. M., & Markley, J. L. (1991). A relational database for sequence-specific protein NMR data. *Journal of Biomolecular NMR, 1*, 217–236.

Shiga, M., Suzuki, K., & Tachikawa, M. (2010). The chemical shift of deprotonated water dimer: Ab initio path integral simulation. *The Journal of Chemical Physics, 132*, Article 114104.

Tereshchenkova, V. F., Goptar, I. A., Kulemzina, I. A., Zhuzhikov, D. P., Serebryakova, M. V., Belozersky, M. A., et al. (2016). Dipeptidyl peptidase 4 – An important digestive peptidase in *Tenebrio molitor* larvae. *Insect Biochemistry and Molecular Biology, 76*, 38–48.

Thiruvengadam, M., Venkidasamy, B., Thirupathi, P., Chung, I. M., & Subramanian, U. (2021). β-Casomorphin: A complete health perspective. *Food Chemistry, 337*, Article 127765.

Vance, J. E., LeBlanc, D. A., & London, R. E. (1997). Cleavage of the X–pro peptide bond by pepsin is specific for the trans isomer. *Biochemistry, 36*, 13232–13240.

Vien, D., Colthup, N. B., Fateley, W. G., & Grasselli, J. G. (1991). *The handbook of infrared and Raman characteristic frequencies of organic molecules*. Amsterdam, The Netherlands: Elsevier.

Wedemeyer, W. J., Welker, E., & Scheraga, H. A. (2002). Proline cis–trans isomerization and protein folding. *Biochemistry, 41*, 14637–14644.

Wüthrich, K. (1986). NMR with proteins and nucleic acids. *Europhysics News, 17*, 11–13.

Supplementary material for "pH-induced changes in β-casomorphin 7 structure studied by ^{1}H nuclear magnetic resonance and Fourier-Transform infrared spectroscopy"

Table S1: Chemical shifts δ (ppm) of synthetic BCM peptides at pH 2.3 and pH 6.7.

Proton	H_N	Hα	Hβ	Hγ	Hδ	Other
pH 2.3						
Isoleucine	8.18	4.10	1.80	1.33	0.75	
	8.25	4.16	1.80	1.12	0.80	
	8.46	4.23	1.80	–	–	
Tyrosine	8.29	4.87	3.15/2.75			7.05
	8.22	4.60	2.79-2.98			6.77
	–	4.33	2.86-3.13			
Phenylalanine	8.17	4.34	3.02/2.90			7.23
	8.16	4.53	2.99/2.88			7.16
	8.04	4.79	2.92.3.06			6.78
Glycine	7.72	4.02/3.92				
	8.02	4.07/3.89				
Proline	–	4.35	2.16	1.87	3.71/3.61	
	–	4.26	1.80	1.81	3.72/3.41	
	–	4.16	1.99	1.73	3.53/3.19	
pH 6.7						
Isoleucine	7.65	4.03	1.68	1.31	0.75	
	7.58	3.97	1.73	1.01	0.80	
	7.89	3.92	1.73	–	–	
Tyrosine	8.29	4.88	3.15/2.75			7.05
	8.19	4.62	2.79-2.98			6.77
	–	4.33	2.86-3.13			
Phenylalanine	8.16	4.32	3.01/2.78			7.23
	8.16	4.52	2.99/2.88			7.16
	8.02	4.79	2.92.3.06			6.78
Glycine	7.75	4.04/3.90				
	8.04	4.11/3.92				
Proline	–	4.17	2.16	1.93–1.74	3.71/3.61	
	–	4.25	1.89	2.02–1.79	3.72/3.41	
	–	4.35	–	2.13–1.83	3.53/3.19	

CHAPTER 4: Temperature- and pH-induced structural changes of β-casomorphin 11-A1 and -A2 studied by Nuclear Magnetic Resonance, Fourier-Transform infrared spectroscopy, chemometrics and molecular modelling

Abstract

The β-casomorphin 11 (BCM11) peptides are embedded in β-casein A1 and A2, and differ by one amino acid at position 8 of the peptide, i.e., His for BCM11 from β-casein A1 (BCM11-A1) and Pro from β-casein A2 (BCM11-A2). This study investigated the structural characteristics of BCM11-A1 and BCM11-A2 at pH 2.5 and 7.0 at 4, 20 or 37°C using Nuclear Magnetic Resonance (NMR), Fourier-Transform infrared (FTIR) spectroscopy, chemometrics and molecular modelling. Both peptides adopt very different secondary and tertiary structure at the observed conditions. The NMR results showed that at pH 2.5, the Tyr^{1}-Pro^{2} and Phe^{3}-Pro^{4} bond in BCM11-A1 had predominantly *cis* isomers and Gly^{5}-Pro^{6} dominates with *cis* isomer, but in BCM11-A2 *trans* isomers were more dominant for all X-Pro bonds. The variation in pH and temperature affected the X-Pro isomerization. At pH 7.0, an increase in temperature led to changes in the secondary structure by reformation of proline-induced turns. The C-terminus of BCM11-A2 forms a compact structure due to bending and formation of proline-induced turns held by hydrogen bonds. BCM11-A1 adopts turn structures and folds into less dense structure with highly exposed hydrophilic parts of the peptide (imidazole ring). Histidine had major impact on the peptide structure in respect to pH and temperature variation that leads to multiple restructuring of the β-turn components. This study brings significant insights in the importance of the amino acids positioning and the structural arrangement of BCM11-A1 and BCM11-A2 that defines their main differences in regards to the genetic variance.

1. Introduction

During enzymatic hydrolysis and/or fermentation, the β-casein (β-CN) polypeptide chain is broken down into smaller peptides (Daniloski *et al*., 2021a; Jinsmaa & Yoshikawa, 1999; Silva & Malcata, 2005). The liberated peptides might influence the functionality of milk and dairy products and possibly affect human health (Sah, Vasiljevic, McKechnie, & Donkor, 2015). One group of peptides that can be released from β-casein are the β-casomorphins (BCMs), all containing the identical N-terminal sequence of Tyr-Pro-Phe (Daniloski, McCarthy & Vasiljevic, 2021b). The nature of BCMs' as μ-receptor agonists that responds to their opioid activity initiated substantial research into opioid activity of the smaller BCMs (3-7 amino acid residues; BCM3-BCM7) (Asledottir *et al*., 2019). Numerous data exist on BCM3-BCM7 liberated from the β-casein (Brantl, Teschemacher, Henschen, & Lottspeich, 1979; Nguyen, Johnson, Busetti & Solah, 2015), but the data on β-casomorphin 11 (BCM11) liberated from β-CN is lacking. Beta-casomorphin 11 (BCM11: β-casein f[60-70]) is a peptide with 11 amino acids (Tyr-Pro-Phe-Pro-Gly-Pro-Ile-X-Asn-Ser-Leu), where X is His or Pro, depending on whether BCM11 is derived from β-casein genetic variant A1 (BCM11-A1) or β-casein variant A2 (BCM11-A2), respectively.

The difference in amino acid at position 8 in the peptide chain of BCM11-A1 and BCM11-A2 could lead to conformational differences, which might affect the release of smaller peptides from BCM11 upon further hydrolysis, e.g., to form BCM7 as a result of the hydrolysis of the Ile^{7}-His^{8} or Ile^{7}-Pro^{8} bond in BCM11-A1 or BCM11-A2, respectively. Asledottir and co-workers (Asledottir *et al*., 2017; Asledottir *et al*., 2018; Asledottir *et al*., 2019) showed that BCM3 and BCM7 were released at lower levels from β-casein A2 than from β-casein A1. It has been hypothesized that the proline rich peptides

would show higher resistance to enzymatic cleavage (Brantl *et al*., 1979). However, some studies also showed that BCM7 can be released from BCM11-A2 as well (Asledottir *et al*., 2019). The differences of the released BCM7 during hydrolysis of BCM11-A1 or BCM11-A2 may arise due to conformational differences and accessibility of the aforementioned Ile-X bond in BCM11.

However, the information on the structure of BCM peptides and their conformational states is scarce. Four interchanging conformational isomers are found in BCM5 (Tyr-Pro-Phe-Pro-Gly) peptide due to *cis-trans* isomerism of Tyr-Pro and Phe-Pro bond confirming stacking of the side chains in the conformational equilibrium (Delaet, Verheyden, Tourwé & Van Binst, 1991). However, for BCM4 (Tyr-Pro-Phe-Pro) it was observed that the peptide does not have clear conformational preferences due to high chemical exchange between the *cis* and *trans* isomers (Goodman & Mierke, 1989). In our previous study, we showed that BCM7 adapts different isomeric conformations that were pH dependent (Markoska, Huppertz, & Vasiljevic, 2021a). The *cis-trans* isomerism and the structure can affect the bond accessibility for enzymes to cleave the peptides (Pal & Chakrabarti, 1999). The protonation state of the side chains of the amino acids was postulated to have a notable impact on structural orientation and functional properties of the peptides (Li & Hong, 2011). In this respect, the tautomerisation of the imidazole ring of His was shown to be dependent on protonation/deprotonation state of the molecule (Li & Hong, 2011). The current study identified the structural changes in the peptides BCM11-A1 and BCM11-A2 as a function of pH (2.5 and 7.0) and temperature (4, 20 and 37 °C). The selected pH was used as intestinal pH in order to confirm structural behaviour of the peptides. The selected temperatures were aiming to accommodate conditions critical for dissociation behaviour of β-CN (4 ^{o}C), room temperature (20 ^{o}C) and human body

temperature (37 °C). The findings facilitate understanding of how the conformational changes may be the reason for differences in the release of BCM peptides from β-casein A1 or A2.

2. Material and Methods

2.1.Sample preparation

Two synthetic 11-amino acid BCM peptides (BCM11-A1 and BCM11-A2) of analytical reagent grade with purity of >98 % were prepared by Mimotopes (Mulgrave, Melbourne, VIC, Australia). The experiments were arranged in a full factorial experimental design with three factors. The first factor is the type of peptide, i.e., BCM11-A1 (H-Tyr-Pro-Phe-Pro-Gly-Pro-Ile-**His**-Asn-Ser-Leu-OH) and BCM11-A2 (H-Tyr-Pro-Phe-Pro-Gly-Pro-Ile-**Pro**-Asn-Ser-Leu-OH). The second factor is pH at two levels, i.e., pH 2.5 and pH 7.0. The third factor is temperature at three levels, i.e., 4, 20 and 37 ºC. The temperatures were selected as significant for storage and dissociation of β-casein from micelle (4 ºC), room temperature (20 ºC) and human body temperature (37 ºC). The experiment results with 12 independent observations that was replicated three times. The samples for the NMR analysis were prepared by dissolving 10 mg of each peptide in 1 mL H_2O:D_2O solution (90:10 %, Sigma-Aldrich, St. Louis, MO, USA). The samples for FTIR analysis were prepared by dissolving the samples in ultra-pure water. The pH of the peptide solutions was adjusted to 2.5 and 7.0 with 0.25 M HCl or 0.25 M NaOH (Sigma-Aldrich, St. Louis, MO, USA), respectively. The pH was measured before each analysis, and adjusted if required. Samples were analysed at temperatures of 4, 20 and 37 ºC. The

temperature control for the NMR analysis was automated by the instrument and for the FTIR analysis was measured with temperature probe.

2.2.Fourier Transform Infrared (FTIR) spectroscopy

A PerkinElmer Frontier FTIR spectrometer (Frontier 1, PerkinElmer, Boston, MA, USA) equipped with a diamond attenuated total reflectance (ATR) cell was used for analysis of the secondary structure of the peptides at different pH and temperature. For this purpose, a spectral range of 4000 to 600 cm^{-1} with a resolution of 4 cm^{-1} and 16 scans was selected for each spectrum. The background spectra were scanned before the start of the measurement for every sample using a blank ATR cell. The spectra of each sample was recorded by refilling the ATR cell. The instrumental conditions were consistent for all sample spectra acquisitions and the background scan. The spectra were processed using Spectragryph (software version 1.2.15, Oberstdorf, Germany). The intense solvent signal coming for the H_2O/D_2O mixture was eliminated from the spectra by subtracting the solvent spectrum. The original spectrum in the region 1700-1000 cm^{-1} was used to detect C=O stretching vibrations in Amide I (1700-160 cm^{-1}), whereas N-H bending and C-N stretching for Amide II and III were detected at 1600-1500 cm^{-1} and 1500-1200 cm^{-1}, respectively (Carbonaro & Nucara, 2010; Curley, Kumosinski, Unruh & Farrell, 1998; Jaiswal *et al*., 2015).

Principal component analysis (PCA) was applied to the second derivative of the FTIR results in the different regions including Amide I, Amide II and Amide III. Origin Pro 2021 software (v. 95 E, OriginLab Corporation, Northampton, MA, USA) was used for PCA data evaluation and reporting. The covariance matrix was used to provide estimate

of the variance in individual variables including pH, temperature and amino acid sequence and further measure whether the selected variables are correlated. PCA provides information by generating sets of principal components (PCs) as coordinated axes with lowest possible loss of information. The PCA analysis was carried out with 95% confidence level.

The absorbance spectra region 1000-1800 cm^{-1} was used for 2D correlation in respect to pH or temperature variation. The averaged spectrum of the pH and temperature block from each peptide was plotted using Origin Pro 2021 software (v. 95 E, OriginLab Corporation, Northampton, MA, USA). The 2D FTIR analysis was performed using synchronous correlation spectra using homo correlation for each peptide in regards to pH and temperature change. The synchronous 2D correlation spectra of FTIR provides information by identifying different peaks and cross peaks showing the in phase variations of the bands in respect to selected variance.

2.3.Nuclear Magnetic Resonance (NMR)

The samples were analysed for primary structure features by NMR using a Bruker Avance spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany) following the previous report (Markoska, Daniloski, Vasiljevic, & Huppertz, 2021b). The instrument was equipped with TXI probe with z-gradient and operated at 600 MHz transmitter frequency. From each sample, 600 µL were transferred into a 5 mm precision NMR tube and inserted in the NMR probe. The instrument conditions were maintained constant for every sample. Before loading the 2D methods, a background spectrum was taken to equilibrate the instrument before applying the new method conditions. The samples were analysed at 4,

20 or 37 ºC using one-dimensional (1D) and two-dimensional (2D) NMR methods. Proton NMR (^{1}H-NMR) spectra was recorded at a spectral width of 9615 Hz and 16 scans in three replications. 2D experiments included correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY). The COSY spectra is phase sensitive with double quantum filter and provides 2D homonuclear shift correlation. The double quantum filter allows anti-phase magnetization of one spin to be converted into double quantum coherence by the second pulse (Rance *et al*., 1983). The TOCSY spectra were phase sensitive, using homonuclear Hartman-Hahn transfer using DIPSI2 sequence for mixing. TOCSY spectra provide information of cross signals that appear between pair of spins which is accomplished by multistep transfer of magnetization across many spins (Ross, Schlotterbeck, Dieterle, & Senn). The NOESY spectra were phase sensitive using 2D homonuclear correlation via dipolar coupling which is due to NOE or chemical exchange. The 2D spectra were recorded at transmitter frequency offset of 2820 Hz and 16 scans. The acquisition mode used was States-TPPI (time-proportional phase incrementation) (Marion & Wüthrich, 1983). The water suppression was achieved using excitation sculpting with gradients allowing for presaturation during relaxation delay in cases of radiation damping.

The spectra were processed using TopSpin (version 4.0.6) software (Bruker BioSpin). The correction of FIDs was performed by 0.3 Hz line-broadening parameter and the phase correction was achieved by 0^{th} and 1^{st} order correction for pk. The baseline correction was performed using qfill fid baseline mode and filter width of 0.1-0.3 ppm. The chemical shift assignment for corresponding amino acids atoms values was performed using sequential assignment and previous values in literature, and BMRB files (Markoska *et al*., 2021a; Seavey, Farr, Westler & Markley, 1991). The temperature dependence of the

chemical shift differences for NH protons of His was fitted by linear model. The R^2 value at pH 2.5 was 0.9997 showing that the chemical shift changes linearly with the temperature change. The R^2 value at pH 7.0 was 0.855 showing non-linear change of the chemical shift with the temperature change.

The sequential assignment of the peptides was performed in the fingerprint region (HN) of the NOESY spectra. NOESY spectra provide information by through space coupling between protons that are in spatial distance less than 5 Å (Wüthrich, 1986). The distance between the protons from the nearby amino acids is shown as sequential distance *JαN(i, i - 1; i - 2, i - 3 . . .*) (Wüthrich, 1986). The spin system of each amino acid was correlated sequentially, by following the *JαN* (*i, i-1, i-2*…) assignment to establish the link between the neighbouring amino acids. The correlation was from Leu^{11} *α*N*i*, then Ser^{10} *α*N*i-1* up to the Pro amino acids. Pro residues break the spin system correlation due to the absence of an HN group in its molecule.

2.4.Molecular modelling

Modelling was carried out utilizing Spartan'16 software (Wavefunction, Inc., 18401 Von Karman Ave., Suite 435, Irvine, CA, USA). The two peptides were constructed using the peptide build facility and the equilibrium conformers were computed using Molecular Mechanics, MMFF. These structures were then subjected to PM3 semi-empirical quantum chemical calculations to establish their final equilibrium geometries. The PM3 semi-empirical quantum chemical calculations are used to optimize the geometry of the molecule and determine the physical and chemical properties. Molecular geometry is the

arrangement of the atoms in the space. The geometry optimization determines the atomic arrangement that makes the molecule stable. In the conformational search of the molecule, the polarities of the chemicals bonds were optimized. The PM3 method was found to be widely useful in conformational studies of different organic, bioorganic molecules (Kumbhar, Kumbhar & Sonawane, 2012; Ustabaş *et al*., 2010).

3. Results and discussion

3.1. Observation of the structure by NMR analysis

The chemical shift of the individual amino acid in the NMR spectra was identified in the COSY and TOCSY spectra. COSY spectra can be used for identification of Hα-protons in the peptides by providing one signal for each amino acid (Griesinger, Otting, Wuethrich & Ernst, 1988). The position of the individual amino acids was detected in the Hα-HN region ($J_{\alpha N}$) of the TOCSY spectra. TOCSY spectra provide information of the neighbouring atom that interact through bond giving a pattern that is specific for each amino acid (Griesinger *et al.,* 1988). The amino acid spin systems of BCM11-A1 and BCM11-A2 analysed at pH 2.5 and 20 ºC showed best spectral resolution and were used as a reference for the changes as a function of changes in temperature and pH. The Hα protons for some amino acids were not identified in the 1D ^{1}H NMR spectrum due to signal's overlapping with the suppressed water signal, extensive exchange with the solvent signal or low signal to noise ratio due to modified temperature and pH conditions. However, the position of most of the chemical shifts was identified using the cross peaks in 2D TOCSY and NOESY spectra. The sequential assignment for *JαN* for both peptides is shown in Supplementary Figure S1.

In the TOCSY spectra for the amino acids preceding Pro residues (X-Pro motifs), two sets of spin systems were observed, with different intensities, and were referred to as *cis* or *trans* isomers of the X-Pro bond (Wedemeyer, Welker & Scheraga, 2002). The isomers of the X-Pro bond were analysed by observing the cross peaks between X(Hα) and Pro(Hδ) for *trans* isomers and X(Hα) and Pro(Hα) for *cis* isomers (Chazin *et al*., 1989). The *cis-trans* isomers were difficult to be distinguished in the peptides with pH 7.0 due to low signal to noise ratio. BCM11-A1 and BCM11-A2 were analysed for *cis-trans* isomers at pH 2.5 and 20 °C for structural differences. The major chemical shift intensities observed in the spectra showed different isomers between BCM11-A1 and BCM11-A2. The presence of NOE cross peaks confirms close proton distance that allows observation of the spins systems for *cis-trans* isomers. The assigned spin systems for both peptides are shown in Figure 1.

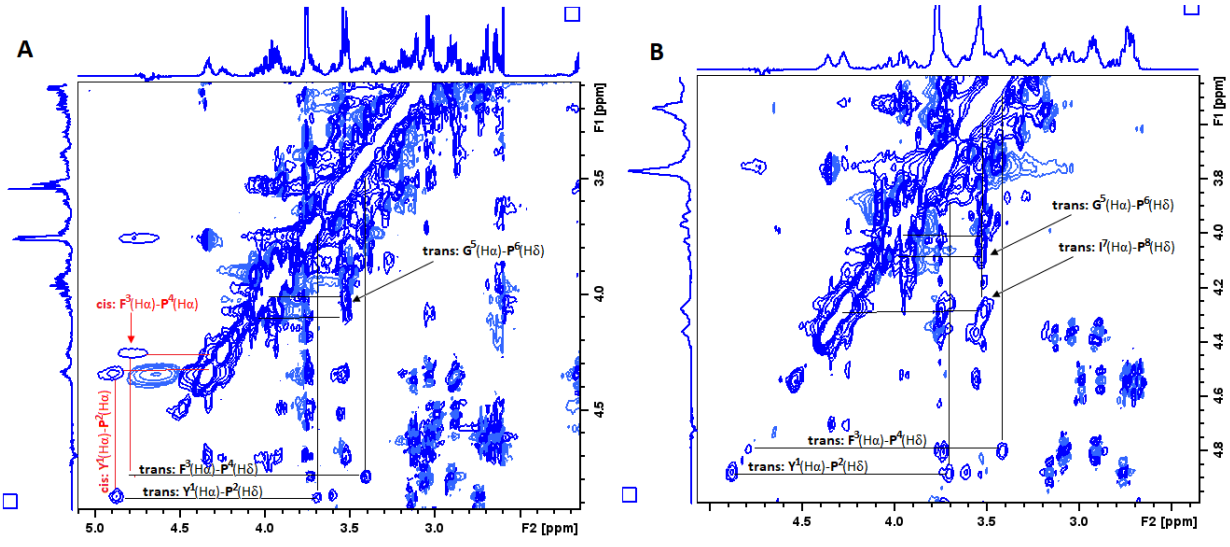


Figure 1**.** NOESY spectra presenting the NOE cross peaks for X-Pro bond from BCM11-A1 (A) and BCM11-A2 (B). Y is Tyr, F is Phe, G is Gly and I is Ile. The spin systems

are observed for *trans* isomers ($d_{\alpha\delta}$) and *cis* isomers ($d_{\alpha\alpha}$). Red colour presents the *cis* isomers and black presents the *trans* isomers.

The X-Pro bonds in the BCM11-A1 adopt both *cis* and *trans* isomers and in BCM11-A2 predominately *trans* isomers. In BCM7 we previously observed presence of both *cis-trans* isomers at all X-Pro bonds except for Gly^5-Pro^6, which was only found in *trans* conformers at pH 6.7 (Markoska *et al*., 2021a). This Gly^5-Pro^6 bond was observed to exist predominately in *trans* conformers for both BCM11-A1 and BCM11-A2 in the current work. In BCM11-A1 Gly^5-Pro^6 showed, only *trans* NOE correlation, and the remaining X-Pro bonds were found to adopt both the *cis* and *trans* conformation. The backbone orientation of the peptide and the chemical shift in the NMR spectra are highly correlated with the ψ dihedral angle of the directly preceding residue (Wishart, Bigam, Holm, Hodges, & Sykes, 1995). The effect is greatly dependent of presence of prolines in the peptide chain. Thus, the ψ distribution is affected due to steric clashes occurring between Hδ of Pro and Hβ of the neighbouring amino acid, which depends on the conditions (MacArthur & Thornton, 1991). Since Gly does not possess Hβ protons, the mentioned steric constraint might not be common for the isomeric rotation of the Gly^5-Pro^6 bond. It is expected that in aqueous solution, peptides to have predominately *trans* conformers due to the lower steric repulsion between the Hα atoms of the X-Pro residues (Kakalis, Kumosinski & Farrell, 1990). However, for the *cis-trans* assignment we considered only the major chemical shift of the spectra, even though the NOESY spectra showed the existence of minor cross peaks that confirms presence of multiple isomers. Therefore, the BCM11 peptides have multiple conformational states that give a highly flexible structural organization. The stability of the *cis-trans* isomers in the peptides is affected by intra- and intermolecular or hydrogen effects in aqueous solution

(Brahmachari, Rapaka, Bhatnagar & Ananthanarayanan, 1982). In aqueous solution, the peptides have high structural flexibility that hinders the precise interpretation of the NOE intensities in terms of sole structure. The presence of *cis-trans* isomers in the peptides confirms peptide bending and formation of turns that can be stabilised by weak connections as hydrogen or Van der Waals bonds.

3.1.1. pH effect on NMR chemical shift

The chemical shifts of the peptides provide information of the backbone dihedral angles, hydrogen bond interaction, and orientation of the aromatic residues (Szilágyi, 1995; Wishart & Nip, 1998). The chemical shifts of BCM11-A1 and BCM11-A2 showed significant pH and temperature dependence. The applied temperature of 4 and 37°C lead to loss of the signal for some of the proton chemical shifts predominately at pH 7.0. Therefore, the pH effect is mainly discussed for the spectra at 20°C to confirm the major variations in the chemical shift of both peptides. The ^{1}H chemical shifts for BCM11-A1 and BCM11-A2 at pH 2.5 and 7.0 are shown in Table 1.

Table 1. ^{1}H chemical shifts for BCM11-A1 and BCM11-A2 at pH 2.5 and 7.0 and 20 °C

^{1}H chemical shifts for BCM11-A1 at pH 2.5 and 7.0 and 20°C (ppm)

Amino acid	pH	HN	Hα	Hβ	Others
Tyr[1]	2.5	8.33	4.87	3.17/2.75	ring: 2,6 Hδ -7.32; 3,5εH -7.24
	7.0	NA	4.86	3.11/2.71	ring: 2,6 Hδ -7.29; 3,5εH-7.20
Pro[2]	2.5		4.32	2.15/1.90	Hγ -2.09; Hδ -3.61/3.32
	7.0		4.32	2.14/1.85	Hγ -2.05; Hδ -3.61/3.31
Phe[3]	2.5	8.07	4.78	3.06/2.94	ring: 2,6 Hδ -6.76; 3,5Hε -7.03; 4Hζ -6.91
	7.0	7.99	4.76	3.06/2.89	ring: 2,6 Hδ -6.75; 3.5Hε -7.03; 4Hζ -6.88
Pro[4]	2.5		4.33	1.98/1.58	Hγ -2.05; Hδ -3.72/3.40
	7.0		4.31	1.88/1.70	Hγ -2.05; Hδ -3.72/3.41
Gly[5]	2.5	7.73	4.03/3.92		
	7.0	7.73	4.00/3.90		
Pro[6]	2.5		4.35	2.10/1.86	Hγ -2.15; Hδ -3.53/3.12
	7.0		4.36	2.16/1.88	Hγ -2.26; Hδ -3.58/3.18
Ile[7]	2.5	8.12	3.99	1.66	Hγ -1.29/1.03; δ1(CH_3) -0.72; δ2(CH_3) -0.74
	7.0	8.11	3.93	1.68	Hγ -1.29/1.25; δ1(CH_3) -1.01; δ2(CH_3) -0.76
His[8]	2.5	8.51	4.61	3.03/3.11	ring: Hε1 -7.19; Hδ2 -7.15; Hδ2 -10.74; Hε2 -10.96
τ	7.0	7.69	4.34	2.95/3.06	ring: Hε1 -7.15; Hδ2 -6.90; Hε1 -9.13
π	7.0	8.10	4.50	2.87/2.96	ring: Hε1 -7.13; Hδ2 -6.93; Hδ1 -9.28
Asn[9]	2.5	8.45	4.60	2.68/2.60	Hδ -7.49/6.73
	7.0	8.13	4.01	2.75/2.65	Hδ -7.47/6.78
Ser[10]	2.5	8.29	4.33	3.75/3.75	
	7.0	8.20	4.32	3.74/3.74	
Leu[11]	2.5	8.18	4.23	1.52/1.52	Hγ -1.36; δH1(CH3) -0.75; δH2(CH3) -0.73
	7.0	7.83	4.08	1.47/1.47	Hγ -1.58; δH1(CH3) -0.75; δH2(CH3) -0.96

^{1}H chemical shifts for BCM11-A2 at pH 2.5 and 7.0 and 20°C (ppm)

Amino acid	pH	HN	Hα	Hβ	Others
Tyr[1]	2.5	8.31	4.86	3.16/2.75	ring: 2,6 Hδ -7.32; 3,5εH -7.24
	7.0	NA	4.86	3.13/2.74	ring: 2,6 Hδ -7.19; 3,5εH-7.25
Pro[2]	2.5		4.39	1.80/2.16	Hγ -1.90; Hδ -3.70/3.63
	7.0		4.33	1.89/1.85	Hγ -1.90; Hδ -3.51/3.31
Phe[3]	2.5	7.98	4.78	3.11/2.94	ring: 2,6 Hδ -6.78; 3,5Hε -7.09; 4Hζ -6.92
	7.0	7.90	4.78	3.07/2.89	ring: 2,6 Hδ -6.75; 3.5Hε -7.04; 4Hζ -6.88
Pro[4]	2.5		4.64	1.78/2.08	Hγ -2.05; Hδ -3.72/3.40
	7.0		4.28	1.75/2.01	Hγ -2.05; Hδ -3.72/3.41
Gly[5]	2.5	7.75	4.05/3.92		
	7.0	7.75	4.05/3.92		
Pro[6]	2.5		4.30	2.30/1.75	Hγ -2.15; Hδ -3.53/3.12
	7.0		4.29	2.24/1.81	Hγ -2.26; Hδ -3.58/3.18
Ile[7]	2.5	8.19	4.28	1.52	Hγ -1.38/1.01; δ1(CH_3) -0.78; δ2(CH_3) -0.83
	7.0	8.19	4.24	1.51	Hγ -1.38/1.36; δ1(CH_3) -1.06; δ2(CH_3) -0.82
Pro[8]	2.5		4.57	1.83/1.94	Hγ -2.28; Hδ -3.78/3.56
	7.0		4.31	1.80/1.90	Hγ -2.02; Hδ -3.74/3.53
Asn[9]	2.5	8.41	4.53	2.69/2.74	Hδ -6.80/7.52
	7.0	8.41	4.52	2.65/2.75	Hδ -6.81/7.51
Ser[10]	2.5	8.11	4.34	3.75/3.75	
	7.0	8.07	4.33	3.73/3.73	
Leu[11]	2.5	8.12	4.29	1.53/1.53	Hγ -1.68; δH1(CH3) -0.76; δH2(CH3) -0.62
	7.0	7.83	4.08	1.46/1.46	Hγ -1.60; δH1(CH3) -0.78; δH2(CH3) -0.81

The HN region showed a stable pH chemical shift for Tyr[1], Gly[5], Ile[7] in both peptides, and this was also observed for Asn[9] in BCM11-A2. The remaining amino acid had upfield shifting of the HN chemical shift that was more intense in BCM11-A1 than in BCM11-A2. His[8] showed two HN chemical shifts at pH 7.0, confirming presence of two conformers. The upfield shift for the HN region of His[8] confirms structural transitions happening in this part of the molecule at pH 7.0. The Hα region showed upfield shifting at pH 7.0, which was present for all Pro residues and Leu[11] for both BCM11-A1 and BCM11-A2 and for Asn[9] and His[8] in BCM11-A1. In BCM11-A1, the Hα upfield shifting of His[8], Asn[9] and Leu[11] confirms that the C-terminal part of this peptide is susceptible to pH-induced structural reformation. The pH effect on this part of BCM11-A1 was observed in the aromatic region (6.5-7.5 ppm) of the NMR spectra (Figure 2A and B).

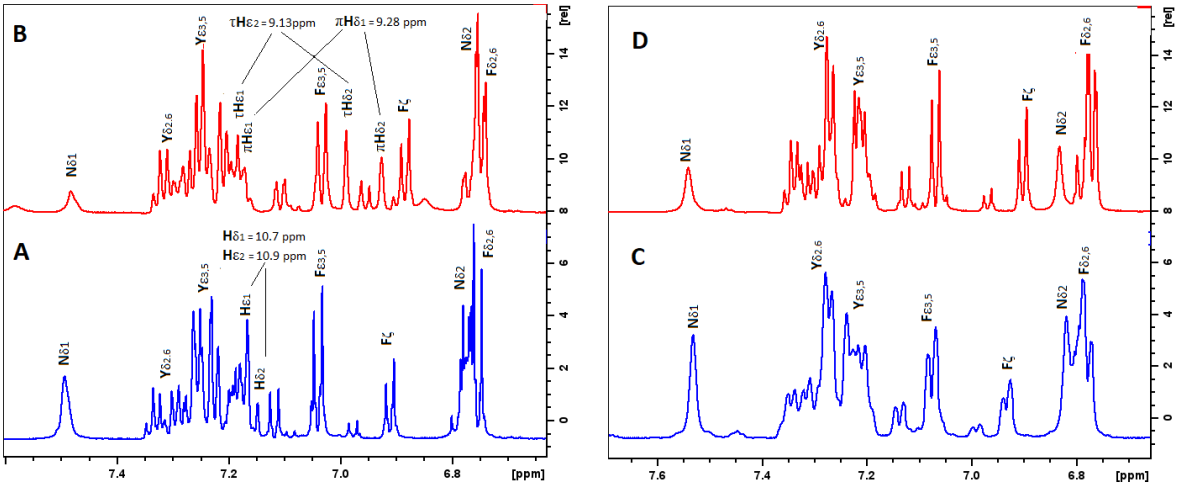


Figure 2**.** Aromatic region of ^{1}H NMR spectra for BCM11-A1 (A,B) and BCM11-A2 (C,D) at pH 2.5 (A,C) and pH 7.0 (B,D) and 20 ºC.

The aromatic region of the ^{1}H NMR spectra of BCM11-A1shown a similar chemical shift for Phe^{3} ring as a result of changes in pH (Figure 2A and B). However, the HN_2 group (HNδ1 and HNδ2) of Asn^{9} at ~7.5 and 6.8 ppm showed lower signal intensity at pH 7.0 due to faster exchange of the ^{1}H with the solvent or possible formation of new hydrogen bonds as proton donor. His^{8} at pH 2.5 had a chemical shift of the Hε1 and Hδ2 from the imidazole ring at ~7.15 ppm (Figure 2A), which in the TOCSY spectra were correlated with cross peaks with Hε2 and Hδ1 at 10.9 and 10.7 ppm, respectively (Table 1). The chemical shifts for Hε2 and Hδ1 were not observed in the ^{1}H NMR spectra because the spectra width was up to 9.0 ppm, but the cross peaks in the TOCSY spectra shown their presence. Thus, the presence of all protons in the imidazole ring of the His^{8} at pH 2.5 confirms that it was fully protonated. The pKa value of the His was observed to be 6.73 for π tautomer and 6.12 for τ tautomer (Tanokura, 1983). Thus, at pH below this pKa the imidazole ring is expected to be more than 50 % protonated that was observed in BCM11-A1 at pH 2.5. At pH 7.0, the Hε1 from the imidazole ring was found in the similar chemical shift (~ 7.15 ppm) as in the peptide at pH 2.5, however Hδ2 showed upfield shifting (Figure 2B). Two sets of chemical shifts for Hε1 and Hδ2 were identified in the aromatic region (Figure 2B) due to presence of different tautomeric structures of imidazole ring of His^{8} (Table 1). In addition, the assigned cross peaks in the TOCSY spectra shown presence of two tautomers in BCM11-A1 at pH 7.0. Therefore, tautomer τ that had deprotonated the Nδ1 and showed only cross peak of Hε1 and HNε2 and π tautomer that had deprotonated the Nε2 and showed only cross peak for Hε1 and HNδ1. The proposed tautomeric structures of the His^{8} in BCM11-A1 are shown in Figure 3.

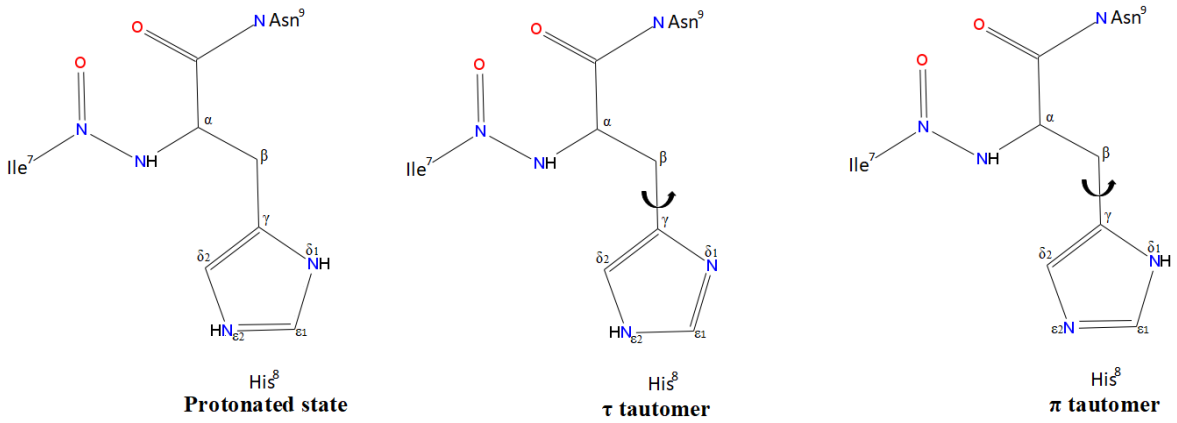


Figure 3**.** Different tautomeric form of imidazole ring from His^8 in BCM11-A1. From left to right: protonated form, τ-tautomer and π-tautomer

The existence of different tautomeric states of imidazole ring of His was observed at pH above 7.0 (Li & Hong, 2011). The deprotonated imidazole ring of His can serve as a proton acceptor and the protonated ring as a proton donor, indicating the ability of His to be involved in hydrogen bonding and structural changes. The hydrogen bonding in BCM11-A1 peptide can occur with the carbonyl from the Leu^{11}, the HN_2 group of the Asn^9, the OH group of the Tyr^1, or the water molecules. The chemical shift of the carbonyl peak appeared at >12 ppm and thus was not observed in our spectra. Moreover, the chemical shift of Tyr^1 ring (Hε3,5) at 7.25 ppm showed merging of two doublets into one larger multiplet when pH was altered from 2.5 to 7.0 (Figure 2A and B) confirming conformational changes of the Tyr^1 ring and possible involvement in new hydrogen bonds. Therefore, the hydrogen bonding happens preferably between the previous observed lower chemical shift of HNδ1 and HNδ2 of Asn^9 as a proton donor, the deprotonated imidazole ring as proton acceptor and the OH group of Tyr^1 as proton

acceptor. The upfield shielding of the Hα and Hβ of Asn^{9} in BCM11-A1 at pH 7.0 (Table 1) show conformational changes occurring for this amino acid. BCM11-A1 at pH 7.0 promotes restructuring and hydrogen bonding that confirms high structural flexibility of this peptide because of pH variation. The NMR analysis of BCM11-A1 shown that His^{8} plays important role in the peptide in regards to the structural changes and involvement in hydrogen bonding. The possible interactions are more favourable at pH 7.0 that plays important role for the peptide bond breakage in the human body and liberalization of smaller opioid peptides. The cleavage of BCMs and the absorption of its fragments happens in the small intestine in the human body, which has a pH of 6.0-7.4 (Fallingborg, 1999). At this pH range, the His from BCM11-A1 is partially deprotonated into different rotameric conformers that gives flexible structural conformations and high accessibility of the Ile^{7}-Pro^{8} bond.

In BCM11-A2, the HN chemical shift had minor pH shielding. The Hα chemical shift had upfield shift at pH 7.0 for all Pro residues and Leu^{11} (Table 1). The upfield chemical shift of the Hα protons indicates the presence of ring current shifts (Dyson *et al*., 1998). The *trans* conformers in BCM11-A2 observed at pH 2.5 were affected at pH 7.0 due to interaction of the rings from the peptide with the backbone. In BCM11-A2, we have two aromatic rings from Tyr^{1} and Phe^{3}. However, only Tyr has a possible binding spot (OH) to interact with another residue from the peptide backbone. To confirm the presence of possible interaction the aromatic region of the NMR spectra was analysed (Figure 2 C and D). In the aromatic region of BCM11-A2, the Hδ1 and Hδ2 from Asn^{9} showed low chemical shift intensity at pH 7.0, confirming involvement of this protons in hydrogen bonding. The aliphatic region of the NMR spectra confirms close proximity of the Hβ of the Asn^{9} with the Hβ or Tyr^{1} confirming close distance of these protons (<5 Å)

(Supplementary Figure 2). The close proximity of the Asn^{9} and Tyr^{1} confirms the possibility of hydrogen bonding. The interaction of the aromatic rings with the Amide protons of BCM11-A2 will promote increase of *cis* population (Dyson *et al*., 1998). The dominant *trans* isomers at pH 2.5 have transited to more *cis* conformers which are known to be favourable in formation of β-turns (Stewart, Sarkar, & Wampler, 1990).

3.1.2. Temperature effect on NMR chemical shifts

The changes in chemical shift of the ^{1}H NMR spectra were analysed in regards to temperature change from 4 to 20 and 37°C. The overlayed spectra of BCM11-A1 and BCM11-A2 are shown in Supplementary Figures S3 and S4, respectively. The chemical shifts of both peptides showed notable temperature shifting. Temperature difference had downfield shifting effect as temperature increased from 4 to 37°C on both peptides regardless the pH. Thus, in reference to spectra at 20°C the peptides at 4°C shifted upfield and peptides at 37°C shifted downfield. The temperature dependent shielding shown to follow similar trend for all observed amino acids except for His^{8}, which showed different shielding in regards to the pH change. The temperature shielding was observed in the TOCSY spectra for both BCM11-A1 and BCM11-A2. An example of the TOCSY spectra presenting the temperature dependent shielding of the proton chemical shift is shown in Supplementary Figure S5.

At both pH values, the HN of BCM11-A1 shifted downfield by ~0.1-0.5 ppm for all observed nuclei in the HN region (7.5 - 9.0 ppm) (Supplementary Figure S3A and B). The HN of the His^{8} had different temperature trend than the HN of the remaining amino acids. The temperature dependence for HN of His^{8} at pH 2.5 and 7.0 is shown in Supplementary Figure S6. At pH 2.5, the shielding effect of HN of His^{8} had similar trend

as the rest of the chemical shift in the spectra, however at pH 7.0 the temperature increase from 4 to 20 and 37°C showed upfield shielding for ~ 0.5 ppm. Thus, the shielding was opposite of the observed downfield shift of the remaining components in the spectra, due to involvement of His^8 in structural reformation of the peptide at pH 7.0 and temperature of 20 and 37°C. The specific temperature transition at pH 7.0 for His^8 is also observed in the aromatic region (Supplementary Figure S3). At pH 7.0, the aromatic region previously shown existence of two tautomeric forms of imidazole ring as observed by two sets of chemical shifts for the deprotonated imidazole ring (Figures 2 and 3). This observation was for BCM11-A1 at 20°C and showed an identical chemical shift for the peptide at 37°C. At 4°C, the peptide was found in a protonated form, which is not expected due to the neutral pH. However, the Ans^9 showed two signals for HNδ1 (7.55 and 7.39 ppm), which were correlated with the HNδ1 and Hδ2 from the imidazole ring (~6.3 and 10.2 ppm) through cross peaks (Supplementary Figure S7). The correlation confirmed that these protons are close enough and interacting through hydrogen exchange, thus the HN of the Asn^9 donates protons to the deprotonated Nδ1 of His^8. Here, we can consider that at 4°C the imidazole ring is involved in hydrogen bonds that keep the peptide into stable conformer. BCM11-A1 showed high structural flexibility of the backbone in respect to temperature and pH variation. Moreover, the unstable imidazole ring gives unsteady conformations to the whole peptide. The minor change in temperature of 4 °C can affect the conformational state of peptides by changes in the *cis-trans* isomerism and intramolecular and intermolecular interactions (Troganis *et al*., 2000). The Hα and aliphatic region at pH 2.5 showed downfield shielding as temperature increased which was consistent for all protons (Supplementary Figure S3). A major change was observed in Hβ of Asn^9 (3.70-3.85 ppm) that showed a low chemical shift intensity when

temperature increased to 37 ^{o}C confirming structural transition happening in C-terminal of the peptide. At pH 7.0 of BCM11-A1 the Hα and aliphatic region shown the same shielding trend as the rest of the spectra. The chemical shift at 0.85-1.20 ppm for all temperature levels showed high intensity in comparison to pH 2.5. This chemical shift appears from the CH_3 of Leu and Ile. At pH 7.0, a large irregular triplet was observed due to coupling of two neighbouring protons that are not chemically equivalent (Supplementary Figure S2B). Moreover, the observed triplet is related to the changes in the adjacent hydrogen atoms of the Ile and Leu side-chain (Wishart *et al*., 1995). The long aliphatic side chains of Ile and Leu can be a key factor for hydrophobic interactions and formation of turns in the peptides (Suh *et al*., 1999).

In BCM11-A2 at pH 2.5, the HN chemical shift of all amino acid shifted downfield for 0.3 ppm as temperature increased (Supplementary Figure S4). The intensity of the chemical shift in the HN region did not show major changes. However, at pH 7.0, the HN downfield shifting with temperature rise was 0.2-0.3 ppm and the major difference was observed in spectra low intensity for the peptide at 37°C, compared to the other temperatures, which is due to faster proton exchange with the solvent. The HN region of BCM11-A2 was more stable to shifting in comparison to that of BCM11-A1. Similarly, the aromatic region (6.5-7.5 ppm) had downfield shielding as temperature increased at both pH values. Major variation was observed in the HNδ1 (7.4, 7.5 and 7.6 ppm) and HNδ2 of Ans^9 (6.7, 6.8 and 6.9 ppm) at 4, 20 and 37°C, respectively. The signal intensity decreased as temperature increase at both pH values. The protons of HN_2 of the Asn^9 can rapidly exchange with the protons from the solvent at minor environmental changes. In addition, the HN_2 can serve as a proton donor for hydrogen bonding. The downfield shifting of the Tyr^1 and Phe^3 rings as temperature rose in the aromatic region confirm

sensitivity to structural reformation. The Hα region and the remaining aliphatic components shifted downfield for approximately 0.2 ppm. The intensity of the Hα of proline in the spectra at 37 °C was low due to loss of the chemical shifts due to suppression of the water signal. The methyl group at 0.8, 1.0 and 1.2 ppm for 4, 20 and 37 °C showed large triplet at pH 7.0 similarly as observed in BCM11-A1.

The shifting of the Hα chemical shift was found to be correlated to the changes in the secondary structure (Shen & Bax, 2007; Wang & Jardetzky, 2002). The upfield shift of Hα chemical shift of up to 0.3 ppm favours α-helix and downfield shifting of up to 0.5 ppm favours β-sheets. For both BCM11-A1 and BCM11-A2, downfield shifting of Hα for 0.2 ppm was observed regardless of pH. The backbone of the peptides involves in conformational reformation which contributing to the downfield shift of the Hα would be assigned to the β-sheet conformers in large molecules as proteins. The size of BCM11 peptides is not sufficient to form β-sheet, but can involve in formation of turns. Prolines in peptides and proteins are known to have β-turn motif that depends on the *cis-trans* isomerism (Che & Marshall, 2006). Thus, the proline ring can adapt different backbone torsional angles depending of its position in regards to its preceding residue. When prolines are found at a *trans* position, the backbone will show extended β-structures and unstable reverse turns, however the *cis* conformer form stable β-turns (Che & Marshall, 2006; Stewart *et al*., 1990). We observed that at pH 2.5 in BCM11-A1, the Tyr^{1}-Pro^{2} and Phe^{3}-Pro^{4} are found in both *cis* and *trans* isomers and for BCM11-A2 the *trans* isomers were predominant for all Pro residues in the peptide backbone. Thus, BCM11-A1 is expected to form more turn like structures and thus more folded structure, and the BCM11-A2 peptide will favour more extended chain structure. However, pH and

temperature variation led to changes in the original isomers and structural reformation of the peptides.

3.2. Observation of the secondary structure components by FTIR

The FTIR results presented the existence of conformational changes of the peptides as a result of pH and temperature variations. The Amide I of the FTIR spectra corresponds to the C=O stretching vibrations of the peptide bond and includes region of 1700-1600 cm^{-1} (Carbonaro & Nucara, 2010). The C-N stretching vibrations and N-H bending are found in Amide II region of the FTIR spectra between 1600 and 1500 cm^{-1} (Curley *et al*., 1998). The Amide III region (1500-1200 cm^{-1}) possesses relatively weak signals that correspond to the H-N bending and C-N stretching vibration of the peptides (Jaiswal *et al*., 2015). The absorbance of Amide I, II and III regions of both peptides at different pH values and temperatures are presented in Figure 4A

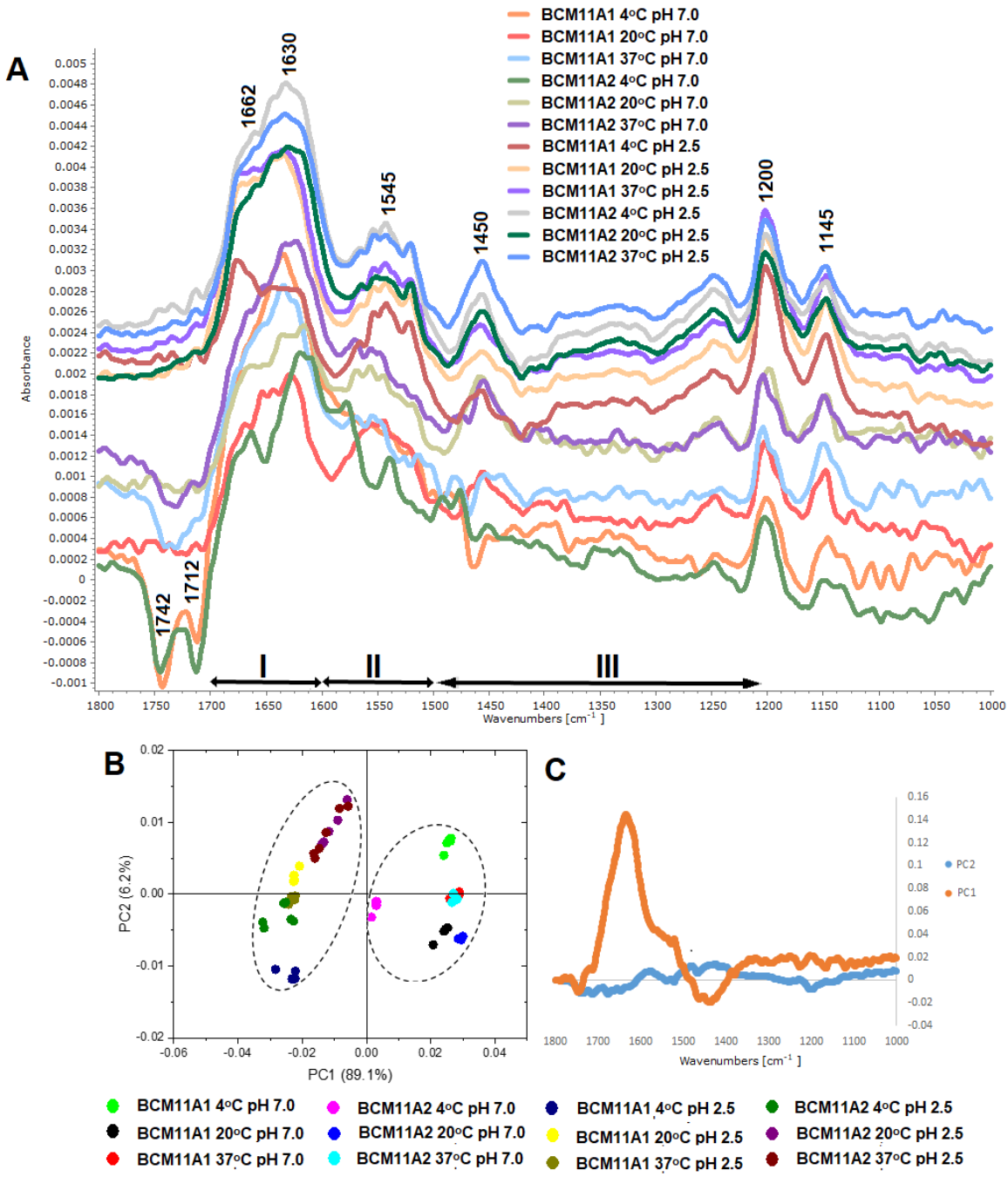


Figure 4. FTIR spectra of BCM11-A1 and BCM11-A2 at pH 2.5 and 7.0 and 4, 20 and 37°C (A). Principal component analysis of BCM11-A1 and BCM11-A2. Loading score (B) and loading plot (C) for FTIR region 1000-1800 cm^{-1}.

The pH had major effect on the spectra distribution in all observed regions. In the Amide I region two major peaks were present (1630 and 1662 cm^{-1}), which confirm presence of turns in the peptides (Vass, Hollósi, Besson, & Buchet, 2003). Both peptides at pH 2.5 had similar peak distribution for the observed peaks in the Amide I region except for BCM11-A1 at 4°C which showed sharp peaks shifted towards higher frequency. The Amide I region for both BCM11-A1 and BCM11-A2 at pH 7.0 had redistribution of the peaks at 1662 and 1630 cm^{-1} that was temperature dependent, confirming to structural reformation. In the Amide II region, the peak with the higher absorbance observed at 1545 cm^{-1} confirming the C-N stretching vibrations and N-H bending of the peptides had only minor changes in peak absorption for peptides at pH 2.5. The peptides at pH 7.0 had peak redistribution in the Amide II region in regards to the temperature change. The peak at 1450 cm^{-1} in the Amide III region appears due to plane bending vibration of the CH_2 groups and stretching vibrations of CN bonds of Pro (Barth, 2000). In the peptides at pH 7.0 the peak at 1450 cm^{-1} followed temperature dependent restructuring. The peak absorbance at 1200 and 1145 cm^{-1} appear to be due to in plane bending vibration of COH from Tyr and vibrations of the CH_2 group of the proline ring, respectively (Barth, 2000). These peaks showed major changes in the absorbance in peptides at pH 7.0 by splitting into two peaks confirming the involvement of hydrogen bonding with the C=O of both peptides (Hienerwadel, Boussac, Breton, Diner, & Berthomieu, 1997). In the FTIR spectra the peptides at pH 7.0 show peaks with negative absorbance at 1712 and 1742 cm^{-}

^{1} that were observed at 4°C and 37°C confirming deprotonation of carboxylic group of the C terminal of the peptide (Davis & Dyer, 2016).

The differences in the peaks were further analysed using the PCA. The PC1 separated the peptides into two groups based on the pH with 89.1 % variance (Figure 4B and C). The loading plot in Figure 4C explains the structural differences of the PC1 and PC2 components based on the PC groupings of the loading score. The PC1 loading plot shows high peak loading at 1635 cm^{-1} for BCM11-A1 and BCM11-A2 at pH 7.0 confirming C=O stretching of the peptide bond due to restructuring of the turns components. The peak at 1630 ±5 cm^{-1} appears due to presence of γ-turns or its repeats (Vass *et al*., 2003). The turns were observed to be more present in peptides at high pH compared to peptides at low pH. Similarly, the peak at 1524 cm^{-1} from the Amide II region showed higher loading for peptides at pH 7.0 confirming higher C-N stretching vibrations and N-H bending of the peptide bonds (Curley *et al*., 1998). For both peptides at pH 2.5, high peak loading was observed at 1430-1460 cm^{-1} due to in plane bending vibration of the CH_2 groups of Pro-ring (Barth, 2000) due to involvement of Pro in structural organization. Another peak loading for both peptides at pH 2.5 was observed at 1746 cm^{-1} which is due presence of COOH of protonated carboxylic acid at low pH (Barth, 2000). The PC2 showed 6.2 % variance without any significant groping of the samples in regards to the observed variables (temperature or pH).

The band components in the FTIR spectra and their correlation with the pH and temperature variation can be analysed using 2D correlation analysis (Figure 5). The presence of off diagonal cross peaks in the 2D synchronous spectrum shows the in phase variations of two bands (Stani, Vaccari, Mitri, & Birarda, 2020). The presence of positive cross peaks confirms that the bands are changing together and the negative cross peaks

confirm that the bands obtain opposite variations in regards to the temperature change (Stani *et al*., 2020). In both BCM11-A1 and BCM11-A2 at pH 2.5 one major auto peak was detected at 1630 cm^{-1} (Amide I) and very weak peaks at 1445, 1200 and 1145 cm^{-1} (Figure 5A and B). The peak intensity for 1630 cm^{-1} in the correlation spectra was more intense for BCM11-A2 at pH 2.5 (Figure 5B). This peak originates from the C=O stretching vibrations of the peptide's bond and also the presence of turns in the molecule (Vass *et al*., 2003).

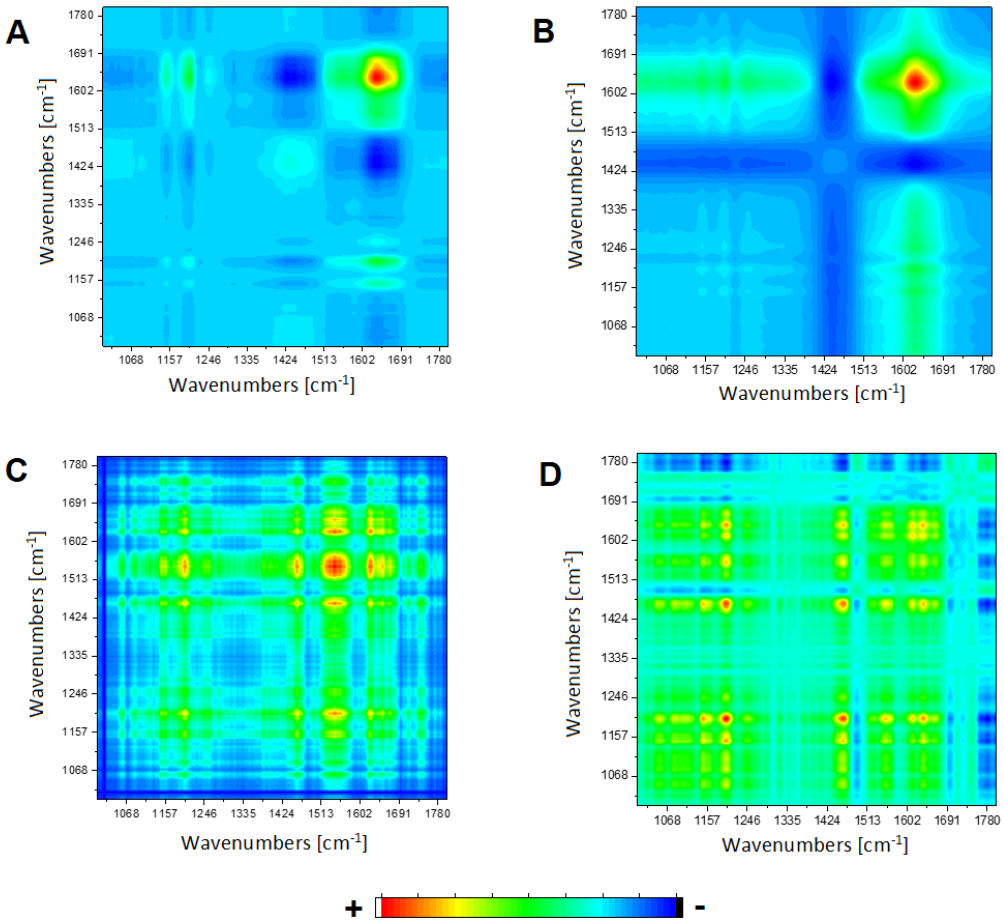


Figure 5**.** Synchronous 2D-correlation plots obtained from FTIR spectra from 1000-1800 cm^{-1} during the temperature ramp of 4, 20 and 37°C for BCM11-A1 at pH 2.5 (A), BCM11-A2 at pH 2.5 (B), BCM11-A1 at pH 7.0 (C) and BCM11-A2 at pH 7.0 (D)

The high loading of this peak for BCM11-A2 can be correlated to the presence of the additional proline in the polypeptide chain which is known as turns indicators (Chou & Fasman, 1977). However, the higher loading of 1445 cm^{-1} for BCM11-A1 is due to vibration stretching from C-H and C-N from His (Barth, 2000). The peak at 1630 cm^{-1} have positive correlation with the minor peaks observed at 1200 and 1145 cm^{-1} meaning that they change together. However, the peaks at 1630 and 1445 cm^{-1} have negative cross

peaks confirming opposite changes with temperature variation. The peaks at 1200 and 1145 cm^{-1} appear due to asymmetric stretching of C-O of the Tyr (Barth, 2000). This component was confirmed to be sensitive to formation of hydrogen bonding with the backbone (Barth, 2000; Hienerwadel *et al*., 1997). The peak at 1450 cm^{-1} appears due to in plane bending vibration of the CH_2 groups of Pro ring (Barth, 2000). For both peptides at pH 2.5 the temperature induced changes of the structure by formation of new hydrogen bonding and turns reformation.

BCM11-A1 and BCM11-A2 at pH 7.0 showed temperature variations by appearance of several peaks observed in different regions of the FTIR spectra (Figure 5C and D). The increased pH had a noticeable effect on the temperature variations on the structure of both peptides. Thus, for BCM11-A1, the major diagonal auto-peaks are observed at 1625, 1543, 1457 and 1200 cm^{-1} (Figure 5C). All four observed diagonal peaks are linked by positive cross peaks confirming that they all change together as a result of temperature variation. For BCM11-A1, the peak at 1545 cm^{-1} shows dominant loading in the 2D spectra (Figure 5C). The peak loading at 1545 cm^{-1} in the Amide II region resulted from N-H bending vibration and C-N stretching vibrations of the peptide which appears from the Pro ring and the imidazole ring of the His (Adochitei & Drochioiu, 2011). The peak loading at 1543 cm^{-1} was with low intensity in the temperature plot of BCM11-A2 (Figure 5D), confirming that N-H vibrations of His are responsible for the high loading in the plot for BCM11-A1 (Figure 5C). In the 2D spectrum of BCM11-A1 at pH 7.0 only positive cross peaks were observed connecting the peaks between Amide I, II and III. Thus, at pH 7.0 the peptide already promoted structure reorganization and is further changed by the temperature variation that is observed by positive correlations between the turns (1625

cm^{-1}) and the C-N and N-H stretching of the prolines and histidine rings and the possible formation of the new hydrogen bonding (peak at 1200 cm^{-1}).

BCM11-A2 at pH 7.0 had less intense peak loading in Amide I region compared to the same peptide at pH 2.5 (Figure 5D). This confirms that there are substantial rearrangement of the structure happening at neutral pH and temperature variations. In the BCM11-A2 at pH 7.0 and temperature variations the observed auto peaks are in Amide I (1630, 1610 cm^{-1}), Amide II (1543 cm^{-1}), Amide III (1445, 1200 and 1145 cm^{-1}). The observed components affected by the temperature variations appeared in the Amide I region were turns (1630 cm^{-1}) and side chains (1610 cm^{-1}). The positive correlation of the peaks in the Amide I region with the weak signal observed in 1543 cm^{-1} shown that the structural components are affected by the C-N and N-H vibrations of the rings of Pro residues. The most intense auto peak at 1445 cm^{-1} appears due to in plane bending vibration of the CH_2 groups of Pro ring and asymmetric stretching of the CH_3 groups of Ile and Leu (Barth, 2000). Thus, the appearing peak at 1610 cm^{-1} can relates to vibrations of the branch chain of the Ile and Leu amino acids in the peptides. Another intense auto peak was detected at 1200 cm^{-1} that previously was observed to be responsible for formation of new hydrogen bonds (Hienerwadel *et al*., 1997). The positive cross peaks appearing in the correlation plot between all the observed components revealed that the structure of BCM11-A2 is changes and new hydrogen bonds are formed because of temperature variation.

3.3. Observation of the tertiary structure components by molecular modelling

The structure of BCM11-A1 and BCM11-A2 was simulated using the modelling program Spartan'16. The program successfully refined the structure of the peptides in the tertiary structure (Figure 6). C_{pk} ovality had index of 2.22 for BCM11-A1 and 2.24 for BCM11-A2 which has been categorised as excellent index for the folding process (Anis, 2008; Hsu, Shu, & Pearn, 2007). Thus, the process capability index is considered poor when the index is <1.0, satisfactory when the index is 1.0-1.5 and excellent when the index is >2.0. The amino acids selected for the model were in protonated form. The folding process was optimized using quantum PM3 calculations performed when the position of all the atoms of the molecular structure were fixed. The optimization process provides the minimum total energy for optimal position of all the atoms in the molecules to be determined. The refined structures for both BCM11-A1 and BCM11-A2 are shown in Figure 6A and B. Both peptide folded into compact structure held by hydrogen bonds. The hydrogen bond donors (HBD) was 10 and 9 and the hydrogen bond acceptor (HBA) was 27 and 26 for BCM11-A1 and BCM11-A2, respectively. Thus, the presence of His makes the BCM11-A1 peptide more in favour of hydrogen interactions than the BCM11-A2 peptide. The model's space filling area (C_{pk} area) and volume (C_{pk} volume) were 1227.6 and 1220.7 for BCM11-A1 and 1218.6 and 1190.5 for BCM11-A2, respectively. Therefore, the BCM11-A2 is more packed into the tertiary structure and BCM11-A1 has structure that is more spacious.

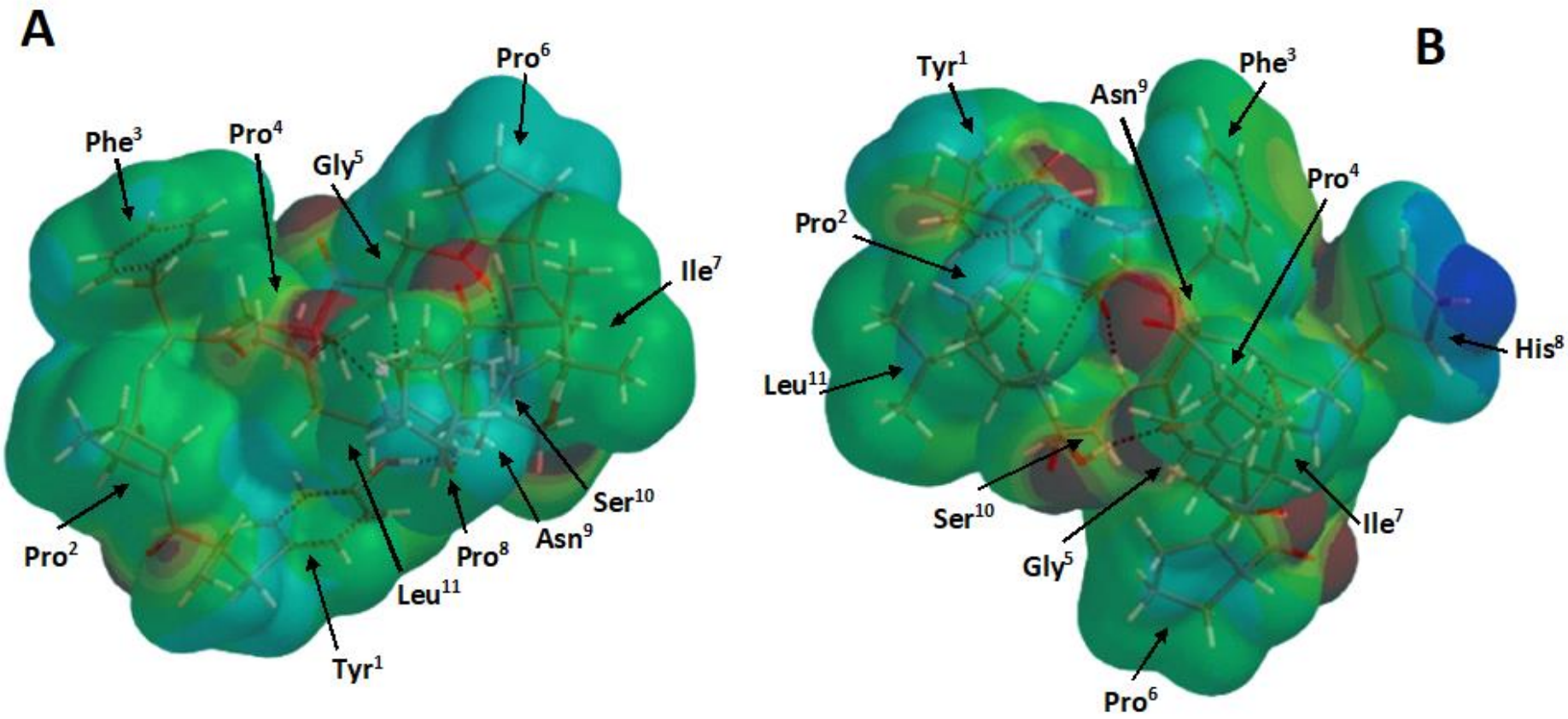


Figure 6**.** Molecular models of BCM11-A2 (A) and BCM11-A1 (B).

In BCM11-A2 (Figure 6A-left) the N-terminal had finely defined turn formed from *cis* isomer of Tyr^{1}- Pro^{2} and stabilized through hydrogen bond between Tyr^{1} ring and NH from Asn^{9}. The Phe^{3}-Pro^{4} bond was in *trans* isomer showing extended chain (top left of the Figure 6A) and the Gly^{5}-Pro^{6} was in *cis* isomer forming second turn in the peptide. The Ile^{7}-Pro^{8} was in *cis* isomer forming third turn that makes the C-terminal of the peptide very compact and stabilized with the hydrogen bonds. The BCM11-A1 peptide had completely different conformational orientation than BCM11-A2 peptide (Figure 6B). All X-Pro bond dominated in *cis* isomer that gives three turns in the peptide. The peptide chain was packed in compact structure held by hydrogen interactions. The Tyr^{1} forms hydrogen bond with the free NH2 group of Asn^{9} similarly as observed in BCM11-A1 peptide. This align with our NMR results confirming the close proximity between the N and C terminal of the peptides and the involvement of Tyr^{1} and Asn^{9} in hydrogen bonding. However, the imidazole ring of the histidine was located at the outer part of the folded peptide, accessible for interactions with the external environment (Figure 6B-right). The imidazole ring in this position is fully protonated and highly hydrophilic (blue). The external position and high hydrophilicity makes the imidazole ring to show strong affinity for water with rapid hydrogen exchangeability. Thus, the Ile^{7}-His^{8} bond is oriented externally, providing accessibility for the enzymes (elastase) to degrade the peptide into smaller fragments (Markoska *et al*., 2021b). The presence of additional Pro in the BCM11-A2 peptide generates higher resistance towards enzymes (Brantl *et al*., 1979). Under selected experimental conditions, the Ile^{7}-Pro^{8} bond is found packed in the

compact structure with the amino acids from the C terminal which makes it less accessible for the enzymes.

In the intestinal digest of both A1 and A2 β-casein variants was observed presence of BCM7 peptides for both genotypes and larger peptides for A2 variant (Lambers, *et al.*, 2021). The heat treatment of milk at 85 °C and 140 °C results with formation of BCM peptides in A1 variant, however in A2 variant the heating restricted the formation of BCM7 peptides (Lambers, *et al.*, 2021). As observed in our current work, the BCM11-A2 peptide shown packed structure that hid the Ile^{7}-Pro^{8} bond internarly and thus restrict the acces for proteolytic reaction and cleveadge to smaler papetide (e.g. BCM7). However, the external exposiure of the Ile^{7}-His^{8} bond can be the reason for the greater amount of BCM7 peptides detected during heat treatment of milk and/or intestinal digestion (Lambers, *et al.*, 2021).

4. Conclusion

The combination of NMR, FTIR and molecular modelling successfully distinguished the secondary and tertiary structure of BCM11-A1 and BCM11-A2. BCM11-A1 was found to have more flexible and loose conformation due to tautomeric behaviour of imidazole ring of His^{8} as a result of pH and temperature variation. BCM11-A2 was observed to have closely packed C terminal end, keeping the Ile^{7}-Pro^{8} buried inside the structure. On the other hand, the Ile^{7}-His^{8} was exposed to the solvent. The imidazole from His^{8} at high pH and temperature variations had rapid proton exchangeability with the water or involve in hydrogen bonding. Thus, this part of the molecule had unstable conformation compared to BCM11-A2 peptide. The *cis-trans* isomerism of the X-Pro bonds shown formation of

multiple turns and peptide folding. At low pH the *trans* isomers dominated in BCM11-A2 and *cis* in BCM11-A1. However, due to high peptide flexibility, the isomers were unstable, that was observed by variations in the secondary structural components especially at high pH and temperature. The study confirmed structural characteristics of BCM11-A1 and BCM11-A2 for the first time. The findings can assist in understanding the conformational preferences of the peptide backbone that can correlate to enzyme affinity for peptide bond cleavage that is found to be the major factor in the debates for the health benefits of A1/A2 milk.

References

Adochitei, A., & Drochioiu, G. (2011). Rapid characterization of peptide secondary structure by FT-IR spectroscopy. *Revue Roumaine de Chimie, 56*(8), 783-791.

Anis, M. Z. (2008). Basic Process Capability Indices: An Expository Review. *International Statistical Review, 76*(3), 347-367.

Asledottir, T., Le, T. T., Petrat-Melin, B., Devold, T. G., Larsen, L. B., & Vegarud, G. E. (2017). Identification of bioactive peptides and quantification of β-casomorphin-7 from bovine β-casein A1, A2 and I after *ex vivo* gastrointestinal digestion. *International Dairy Journal, 71*, 98-106.

Asledottir, T., Le, T. T., Poulsen, N. A., Devold, T. G., Larsen, L. B., & Vegarud, G. E. (2018). Release of β-casomorphin-7 from bovine milk of different β-casein variants after *ex vivo* gastrointestinal digestion. *International Dairy Journal, 81*, 8-11.

Asledottir, T., Picariello, G., Mamone, G., Ferranti, P., Røseth, A., Devold, T. G., & Vegarud, G. E. (2019). Degradation of β-casomorphin-7 through *in vitro* gastrointestinal and jejunal brush border membrane digestion. *Journal of Dairy Science, 102*(10), 8622-8629.

Barth, A. (2000). The infrared absorption of amino acid side chains. *Progress in Biophysics and Molecular Biology, 74*(3), 141-173.

Brahmachari, S. K., Rapaka, R. S., Bhatnagar, R. S., & Ananthanarayanan, V. S. (1982). Proline-containing β-turns in peptides and proteins. II. Physicochemical studies on tripeptides with the Pro-Gly sequence. *Biopolymers, 21*(6), 1107-1125.

Brantl, V., Teschemacher, H., Henschen, A., & Lottspeich, F. (1979). Novel Opioid Peptides Derived from Casein (β-Casomorphins). I. Isolation from Bovine Casein Peptone. *360*(2), 1211-1224.

Carbonaro, M., & Nucara, A. (2010). Secondary structure of food proteins by Fourier Transform spectroscopy in the mid-infrared region. *Amino Acids, 38*(3), 679-690.

Chazin, W. J., Kördel, J., Drakenberg, T., Thulin, E., Brodin, P., Grundström, T., & Forsén, S. (1989). Proline isomerism leads to multiple folded conformations of calbindin D9k: direct evidence from two-dimensional 1H NMR spectroscopy. *Proceedings of the National Academy of Sciences, 86*(7), 2195-2198.

Che, Y., & Marshall, G. R. (2006). Impact of *Cis*-proline analogs on peptide conformation. *Biopolymers, 81*(5), 392-406.

Chou, P. Y., & Fasman, G. D. (1977). β-turns in proteins. *Journal of Molecular Biology, 115*(2), 135-175.

Curley, D. M., Kumosinski, T. F., Unruh, J. J., & Farrell, H. M. (1998). Changes in the Secondary Structure of Bovine Casein by Fourier Transform Infrared Spectroscopy: Effects of Calcium and Temperature1. *Journal of Dairy Science, 81*(12), 3154-3162.

Daniloski, D., Cunha, N. M. D., McCarthy, N. A., O'Callaghan, T. F., McParland, S., & Vasiljevic, T. (2021a). Health-related outcomes of genetic polymorphism of bovine β-casein variants: A systematic review of randomised controlled trials. *Trends in Food Science & Technology, 111*, 233-248.

Daniloski, D., McCarthy, N. A., & Vasiljevic, T. (2021b). Bovine β-Casomorphins: Friends or Foes? A comprehensive assessment of evidence from *in vitro* and *ex vivo* studies. *Trends in Food Science & Technology, 116*, 681-700.

Davis, C. M., & Dyer, R. B. (2016). The role of electrostatic interactions in folding of β-proteins. *Journal of the American Chemical Society, 138*(4), 1456-1464.

Delaet, N. G. J., Verheyden, P. M. F., Tourwé, D., & Van Binst, G. (1991). Assignment of the 1H-nmr resonances of the four rotamers of β-casomorphin-5 in DMSO. *Biopolymers, 31*(12), 1409-1416.

Dyson, H. J., Bolinger, L., Feher, V. A., Osterhout Jr, J. J., Yao, J., & Wright, P. E. (1998). Sequence requirements for stabilization of a peptide reverse turn in water solution. *European Journal of Biochemistry, 255*(2), 462-471.

Fallingborg, J. (1999). Intraluminal pH of the human gastrointestinal tract. *Danish Medical Bulletin, 46*(3), 183-196.

Goodman, M., & Mierke, D. F. (1989). Configurations of morphiceptins by proton and carbon-13 NMR spectroscopy. *Journal of the American Chemical Society, 111*(10), 3489-3496.

Griesinger, C., Otting, G., Wuethrich, K., & Ernst, R. R. (1988). Clean TOCSY for proton spin system identification in macromolecules. *Journal of the American Chemical Society, 110*(23), 7870-7872.

Hienerwadel, R., Boussac, A., Breton, J., Diner, B. A., & Berthomieu, C. (1997). Fourier Transform Infrared Difference Spectroscopy of Photosystem II Tyrosine D Using Site-Directed Mutagenesis and Specific Isotope Labeling. *Biochemistry, 36*(48), 14712-14723.

Hsu, B. M., Shu, M. H., & Pearn, W. L. (2007). Measuring process capability based on Cpmk with gauge measurement errors. *Quality and Reliability Engineering International, 23*(5), 597-614.

Jaiswal, P., Jha, S. N., Borah, A., Gautam, A., Grewal, M. K., & Jindal, G. (2015). Detection and quantification of soymilk in cow–buffalo milk using Attenuated Total Reflectance Fourier Transform Infrared spectroscopy (ATR–FTIR). *Food Chemistry, 168*, 41-47.

Jinsmaa, Y., & Yoshikawa, M. (1999). Enzymatic release of neocasomorphin and β-casomorphin from bovine β-casein. *Peptides, 20*(8), 957-962.

Kakalis, L. T., Kumosinski, T. F., & Farrell, H. M. (1990). A multinuclear, high-resolution NMR study of bovine casein micelles and submicelles. *Biophysical Chemistry, 38*(1), 87-98.

Kumbhar, N. M., Kumbhar, B. V., & Sonawane, K. D. (2012). Structural significance of hypermodified nucleic acid base hydroxywybutine (OHyW) which occur at 37th position in the anticodon loop of yeast tRNAPhe. *Journal of Molecular Graphics and Modelling, 38*, 174-185.

Lambers, T.T., Broeren, S., Heck, J., Bragt, M. and Huppertz, T., 2021. Processing affects beta-casomorphin peptide formation during simulated gastrointestinal digestion in both A1 and A2 milk. *International Dairy Journal*, *121*, p.105099.

Li, S., & Hong, M. (2011). Protonation, Tautomerization, and Rotameric Structure of Histidine: A Comprehensive Study by Magic-Angle-Spinning Solid-State NMR. *Journal of the American Chemical Society, 133*(5), 1534-1544.

MacArthur, M. W., & Thornton, J. M. (1991). Influence of proline residues on protein conformation. *Journal of Molecular Biology, 218*(2), 397-412.

Marion, D., & Wüthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1H-1H spin-spin coupling constants in proteins. *Biochemical and Biophysical Research Communications, 113*(3), 967-974.

Markoska, T., Daniloski, D., Vasiljevic, T., & Huppertz, T. (2021b). Structural changes of β-casein induced by temperature and ph analysed by nuclear magnetic resonance, Fourier-Transform infrared spectroscopy, and chemometrics. *Molecules, 26*(24), 7650.

Markoska, T., Huppertz, T., & Vasiljevic, T. (2021a). pH-induced changes in β-casomorphin 7 structure studied by 1H nuclear magnetic resonance and Fourier-Transform infrared spectroscopy. *International Dairy Journal, 121*, 105106.

Nguyen, D. D., Johnson, S. K., Busetti, F., & Solah, V. A. (2015). Formation and Degradation of Beta-casomorphins in Dairy Processing. *Critical Reviews in Food Science and Nutrition, 55*(14), 1955-1967.

Pal, D., & Chakrabarti, P. (1999). *Cis* peptide bonds in proteins: residues involved, their conformations, interactions and locations11Edited by J. M. Thornton. *Journal of Molecular Biology, 294*(1), 271-288.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983). Improved spectral resolution in COSY 1H NMR spectra of proteins via double quantum filtering. *Biochemical and Biophysical Research Communications, 117*(2), 479-485.

Ross, A., Schlotterbeck, G., Dieterle, F., & Senn, H. Chapter 3-NMR Spectroscopy Techniques for Application to 280 Metabonomics. *The Handbook of Metabonomics and Metabolomics, edited by Lindon, JC, Nicholson, JK, and Holmes, E*, 55-112.

Sah, B. N. P., Vasiljevic, T., McKechnie, S., & Donkor, O. N. (2015). Identification of Anticancer Peptides from Bovine Milk Proteins and Their Potential Roles in Management of Cancer: A Critical Review. *Comprehensive Reviews in Food Science and Food Safety, 14*(2), 123-138.

Seavey, B. R., Farr, E. A., Westler, W. M., & Markley, J. L. (1991). A relational database for sequence-specific protein NMR data. *Journal of Biomolecular NMR, 1*(3), 217-236.

Shen, Y., & Bax, A. (2007). Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *Journal of Biomolecular NMR, 38*(4), 289-302.

Silva, S. V., & Malcata, F. X. (2005). Caseins as source of bioactive peptides. *International Dairy Journal, 15*(1), 1-15.

Stani, C., Vaccari, L., Mitri, E., & Birarda, G. (2020). FTIR investigation of the secondary structure of type I collagen: New insight into the amide III band. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 229*, 118006.

Stewart, D. E., Sarkar, A., & Wampler, J. E. (1990). Occurrence and role of*cis* peptide bonds in protein structures. *Journal of Molecular Biology, 214*(1), 253-260.

Suh, J.-Y., Lee, Y.-T., Park, C.-B., Lee, K.-H., Kim, S.-C., & Choi, B.-S. (1999). Structural and functional implications of a proline residue in the antimicrobial peptide gaegurin. *European Journal of Biochemistry, 266*(2), 665-674.

Szilágyi, L. (1995). Chemical shifts in proteins come of age. *Progress in Nuclear Magnetic Resonance Spectroscopy, 27*(4), 325-442.

Tanokura, M. (1983). 1H-NMR study on the tautomerism of the imidazole ring of histidine residues: I. Microscopic pK values and molar ratios of tautomers in histidine-containing peptides. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 742*(3), 576-585.

Supplementary material for: Temperature and pH-induced structural changes of β-casomorphin 11-A1 and -A2 studied by Nuclear Magnetic Resonance, Fourier-Transform infrared spectroscopy, chemometrics and molecular modelling

Tatijana Markoska[a], Davor Daniloski[a,b], Todor Vasiljevic[a], John Orbell[a] and Thom Huppertz[a,c,d]

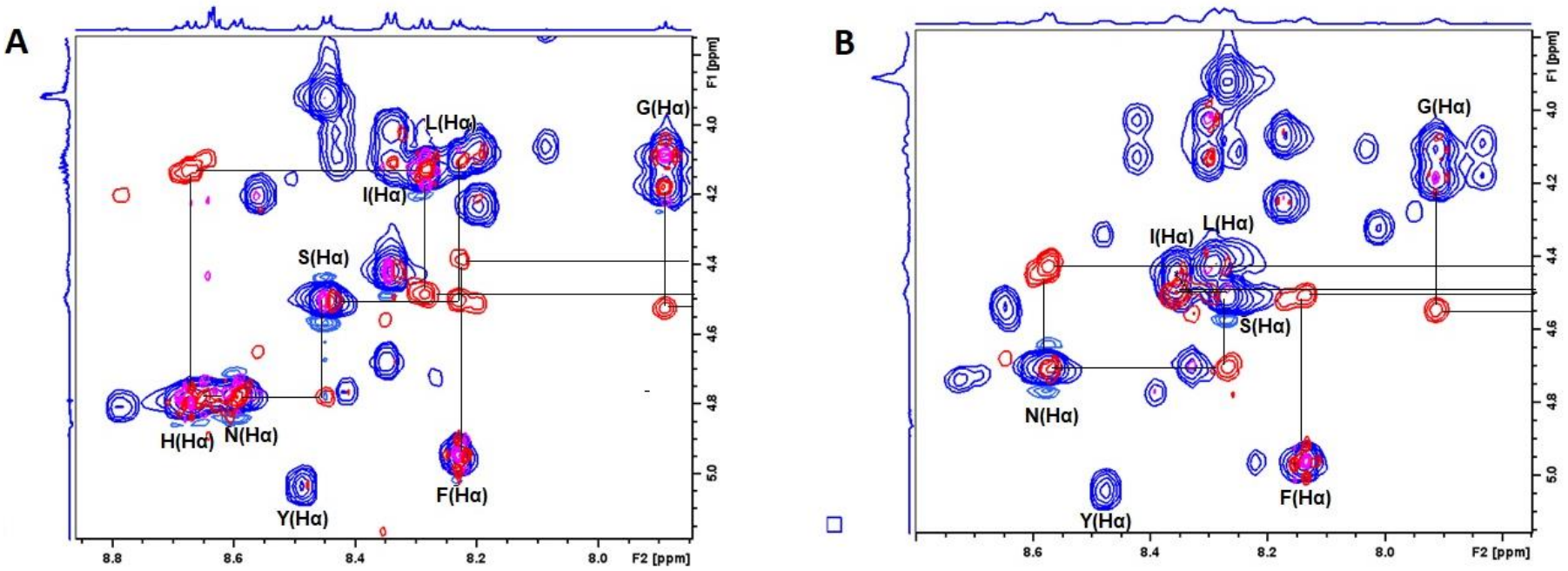


Figure S1. Sequential assignment of the NH-H_α region of the amino acids in BCM11-A1 (A) and BCM11-A2 (B) peptides at pH 2.5 at temperature of 20°C. Red peaks are NOESY spectra, blue peaks are TOCSY spectra.

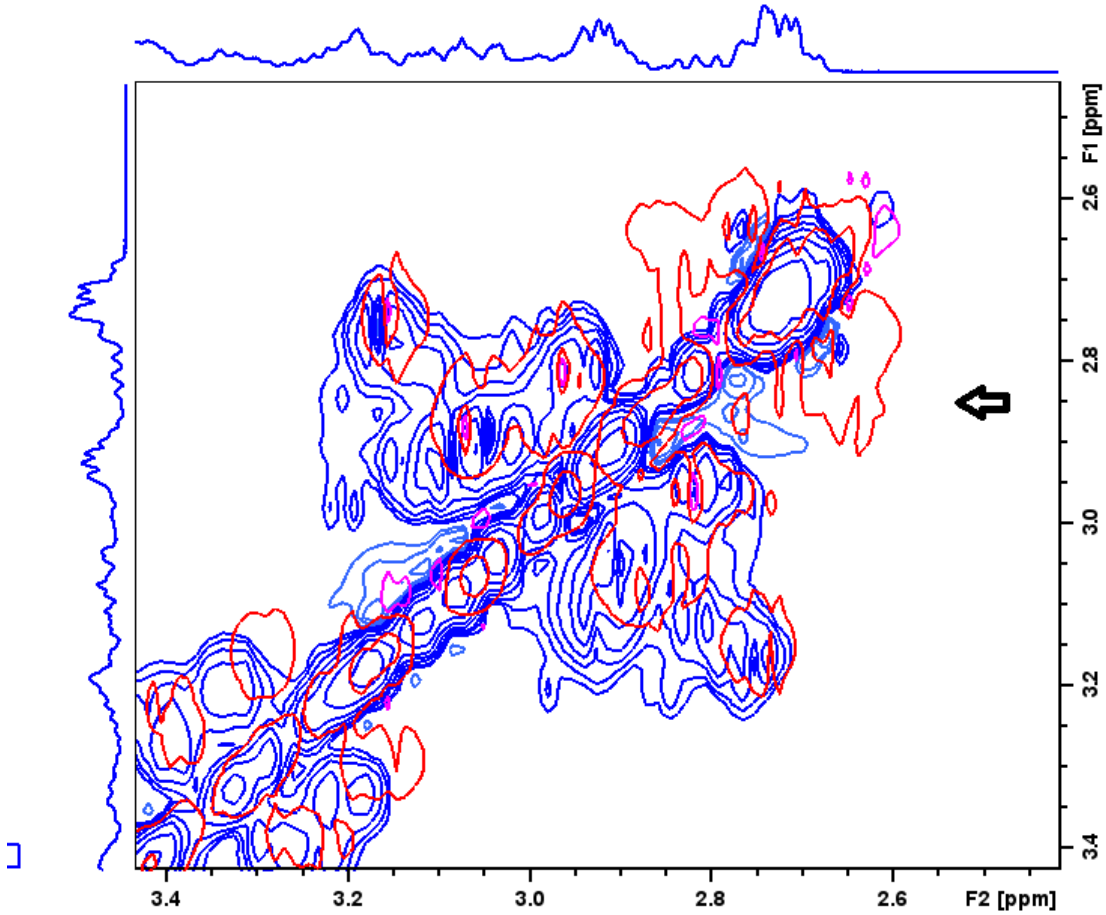


Figure S2 NOESY spectra of BCM11-A2 at pH 2.5 (blue) and BCM11-A2 at pH 7.0 (red). The cross peak observed at ~2.9 ppm correlate Hβ from Asn^{9} with Hβ from Tyr^{1}.

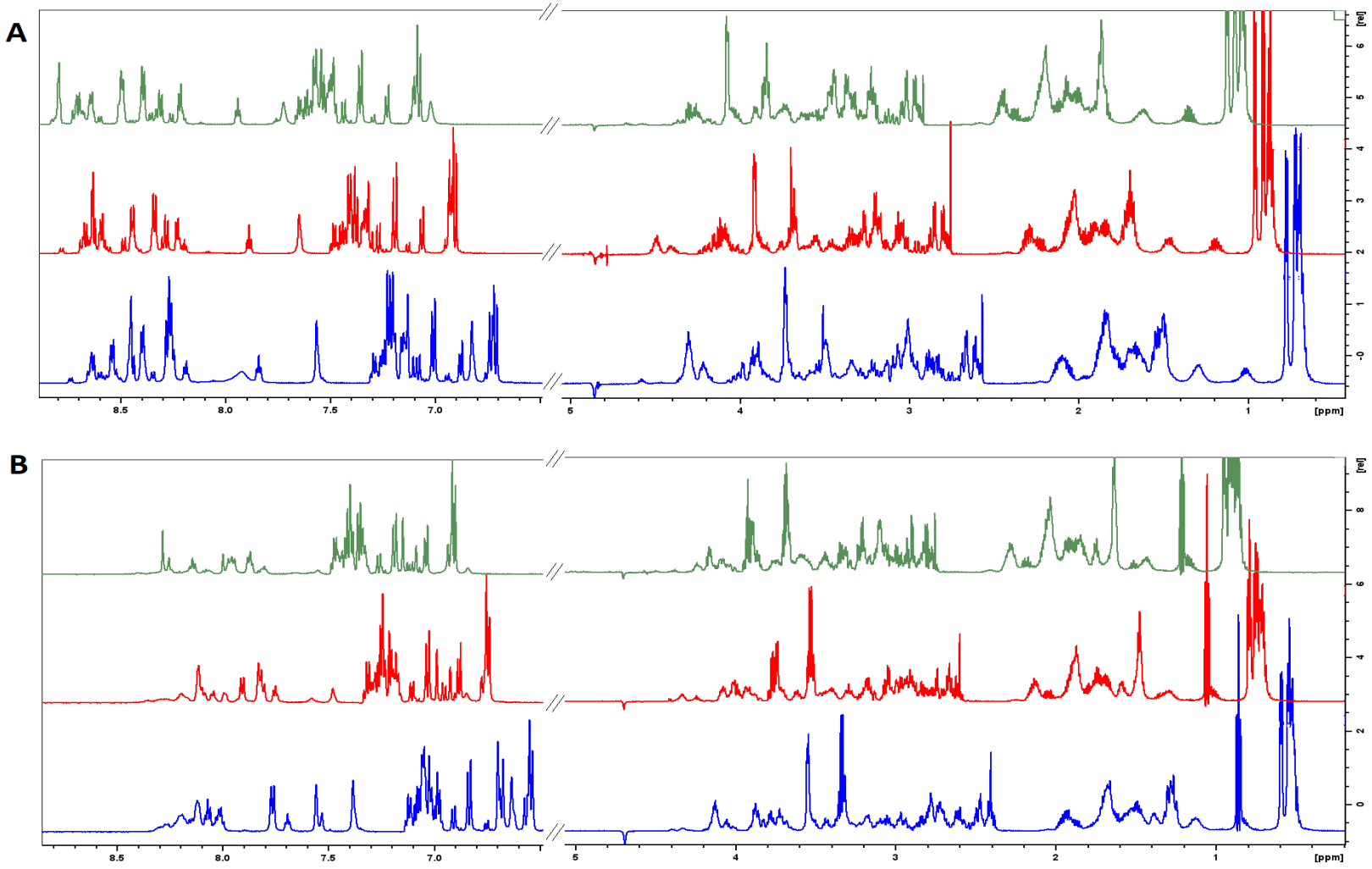


Figure S3: ^{1}H NMR spectra of BCM11-A1 at pH 2.5 (A), and pH 7.0 (B) at 4 (blue), 20 (red) or 37 (green)°C.

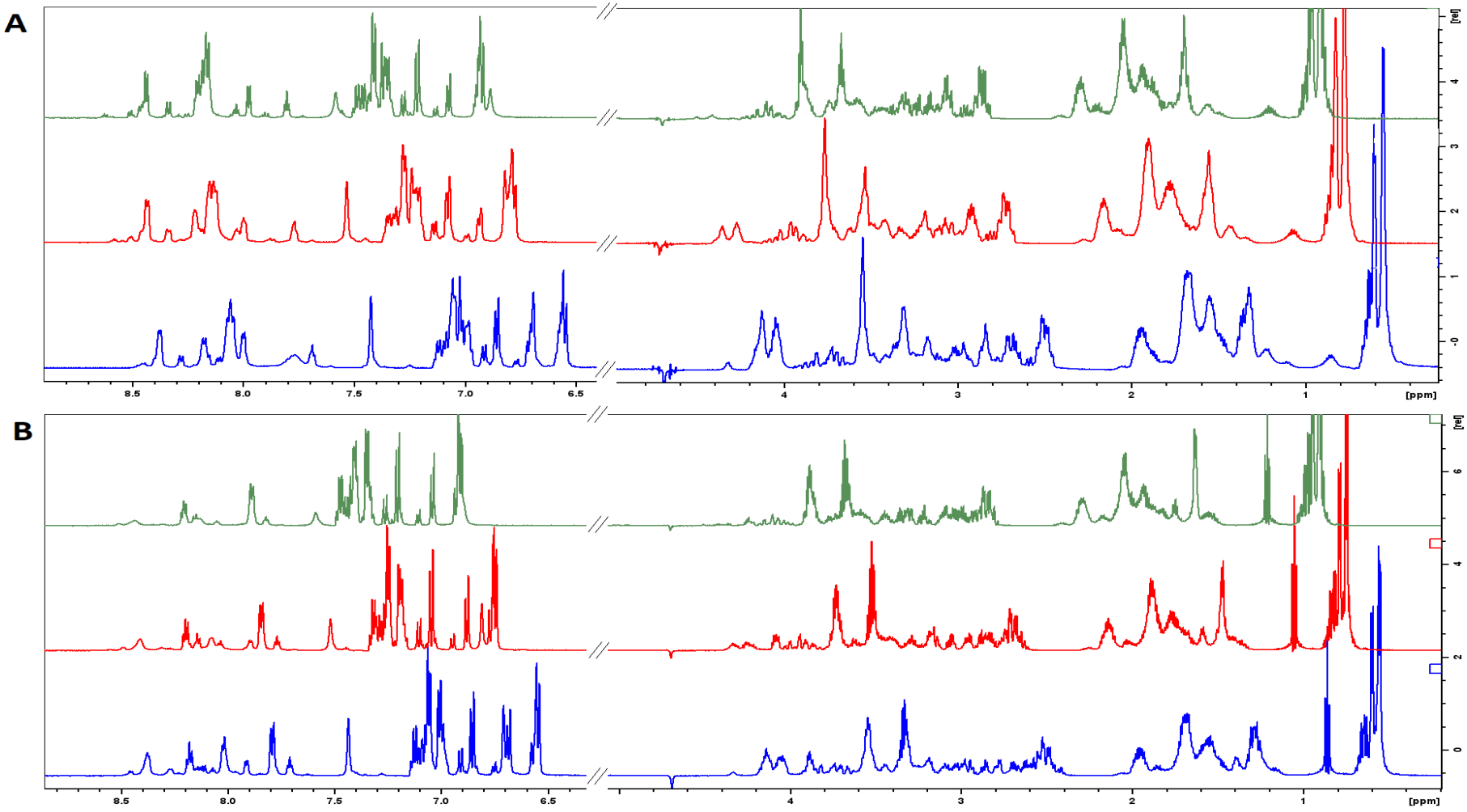


Figure S4: ^{1}H NMR spectra of BCM11-A2 at pH 2.5 (A) and pH 7.0 (B) at 4 (blue), 20 (red) and 37 (green) ºC.

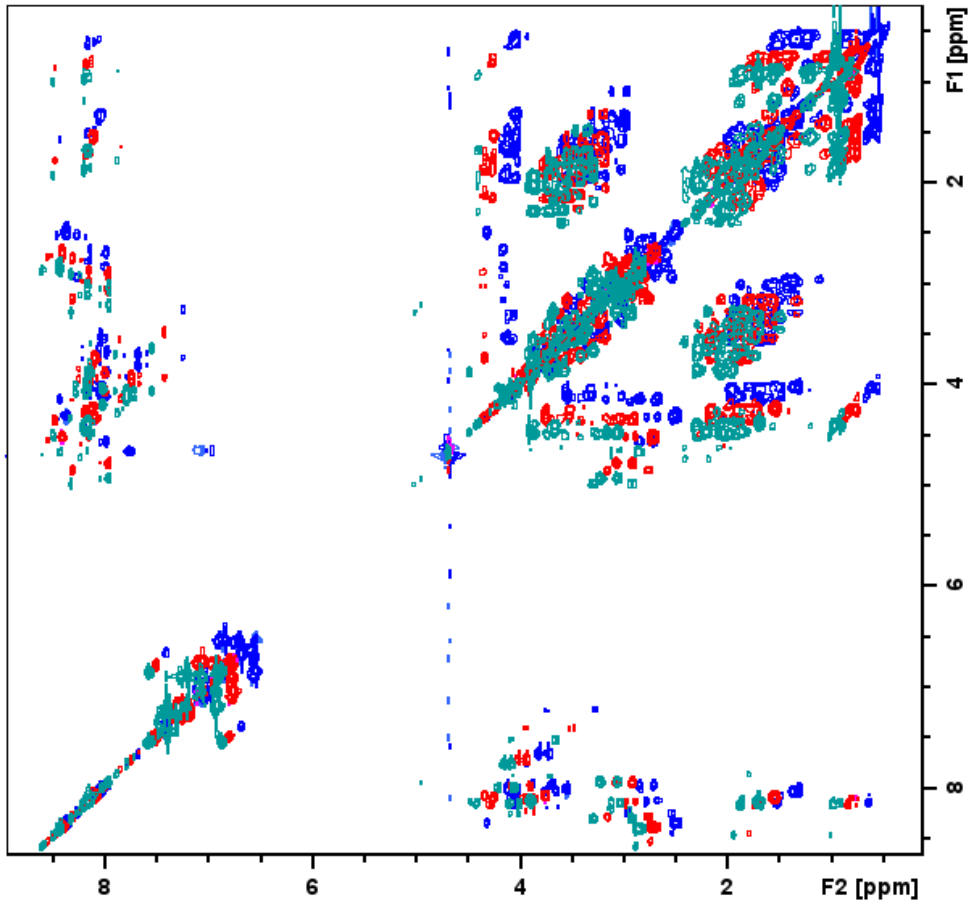


Figure S5. TOCSY spectra presenting temperature dependent chemical shift of BCM11 peptides at 4 °C (blue). 20 °C (red) and 37 °C (green).

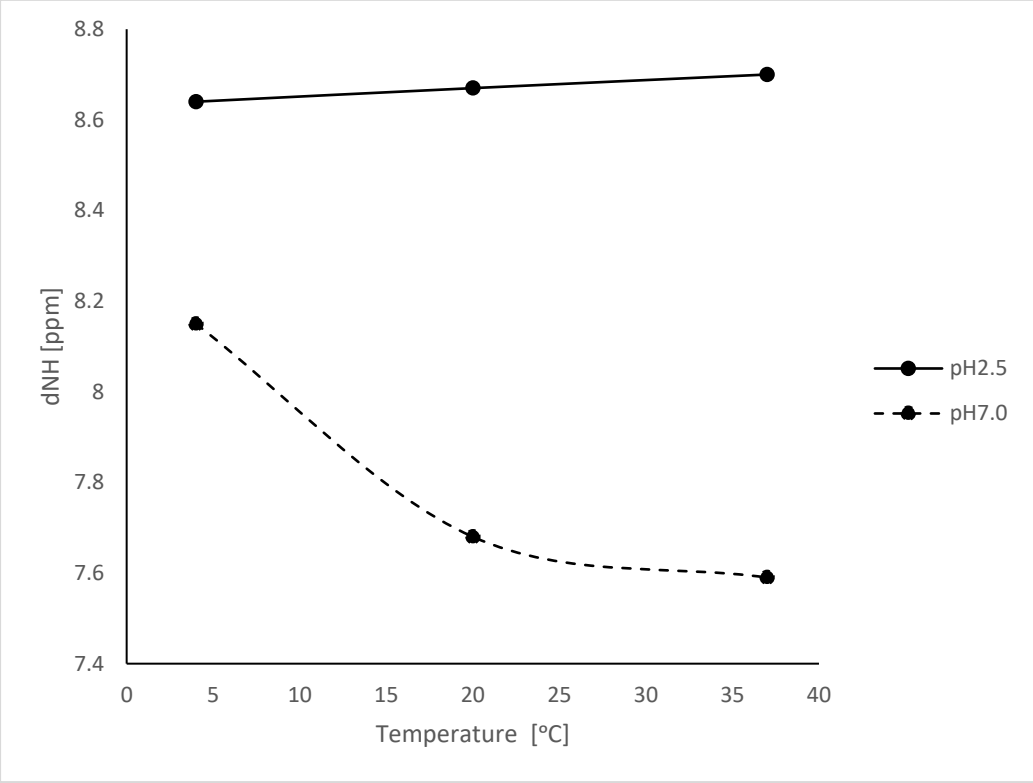


Figure S6. Temperature dependence of dNH of His^{8} in BCM11-A1 at pH 2.5 and pH 7.0.

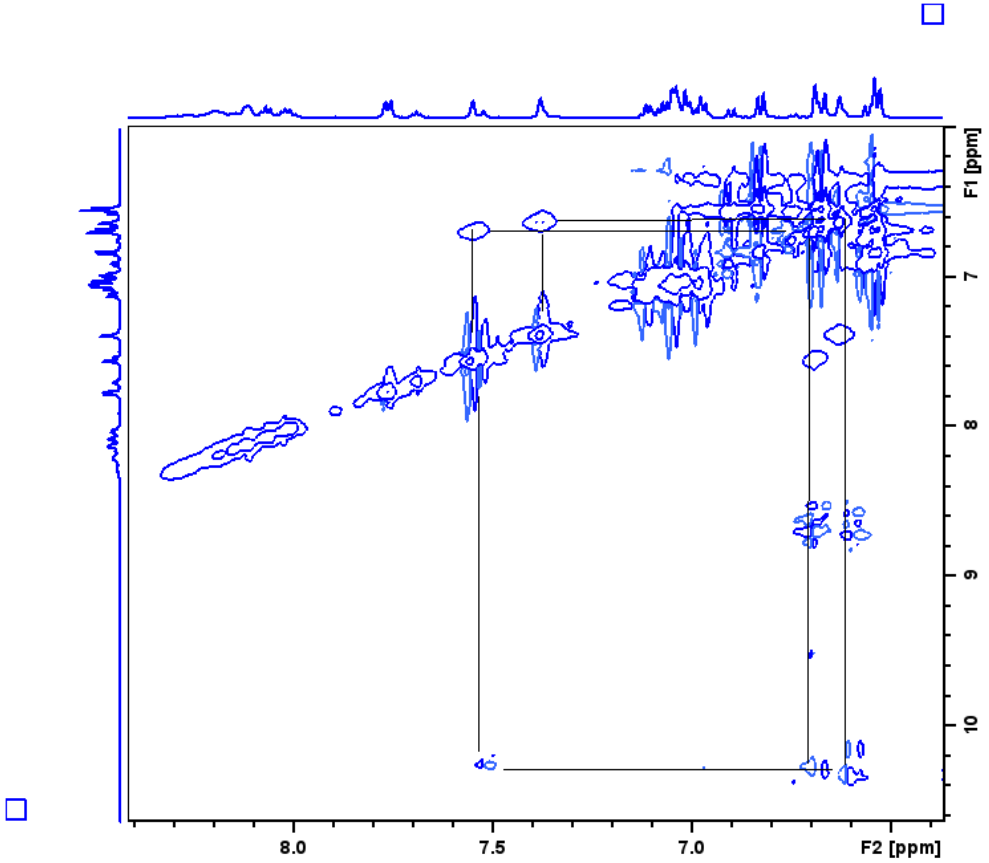


Figure S7. TOCSY correlation for His^{8}- Asn^{9}

CHAPTER 5: Structural changes of β-casein induced by temperature and pH analysed by Nuclear Magnetic Resonance, Fourier-Transform Infrared spectroscopy, and chemometrics

DocuSign Envelope ID: C7AFBA28-8EE3-42F6-BA0C-B4D1B5ADAD78

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book: Structural Changes of β-Casein Induced by Temperature and pH Analysed by Nuclear Magnetic Resonance, Fourier-Transform Infrared Spectroscopy, and Chemometrics

Surname: Markoska First name: Tatijana

Institute: Institute for Sustainable Industries and Liveat Candidate's Contribution (%): 80

Status:

Accepted and in press: ✔ Date: 17/12/2021

Published: ✔ Date: 17/12/2021

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Tatijana Markoska Digitally signed by Tatijana Markoska Date: 2023.03.05 18:50:48 +11'00' | 05/03/2023

Signature **Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83 776 954 731
CRICOS Provider No. 00124K (Melbourne)
02475D (Sydney), RTO 3113

DocuSign Envelope ID: C7AFBA28-8EE3-42F6-BA0C-B4D1B5ADAD78

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

College of Health and Biomedicine, Victoria University, Werribee Campus, Melbourne, Victoria, Australia.

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Tatijana Markoska	80	Experiment design and execution,data analysis, manuscript drafting, revision	[redacted]	30/3/23
Davor Daniloski	4	Concept development, manuscript editing and revision		06/04/23
Todor Vasiljevic	8	Concept development and manuscript editing and reviewing		30/03/2023
Thom Huppertz	8	Designing experiments, manuscript editing and revision and journal submission		3/30/2023

Updated: September 2019

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83 776954 731
CRICOS Provider No. 00124K (Melbourne)
02475D (Sydney) RTO 3113

molecules

MDPI

Article

Structural Changes of β-Casein Induced by Temperature and pH Analysed by Nuclear Magnetic Resonance, Fourier-Transform Infrared Spectroscopy, and Chemometrics

Tatijana Markoska [1], Davor Daniloski [1,2], Todor Vasiljevic [1] and Thom Huppertz [1,3,4,*]

[1] Advanced Food Systems Research Unit, Institute for Sustainable Industries & Liveable Cities, College of Health and Biomedicine, Victoria University, Melbourne, VIC 8001, Australia; tatijana.markoska@live.vu.edu.au (T.M.); davor.daniloski@live.vu.edu.au (D.D.); todor.vasiljevic@vu.edu.au (T.V.)
[2] Teagasc Food Research Centre, Food Chemistry and Technology Department, Moorepark, Fermoy, P61 C996 Cork, Ireland
[3] FrieslandCampina, 3818 LE Amersfoort, The Netherlands
[4] Food Quality and Design Group, Wageningen University & Research, 6808 WG Wageningen, The Netherlands
* Correspondence: thom.huppertz@wur.nl

Abstract: This study investigated structural changes in β-casein as a function of temperature (4 and 20 °C) and pH (5.9 and 7.0). For this purpose, nuclear magnetic resonance (NMR) and Fourier-transform infrared (FTIR) spectroscopy were used, in conjunction with chemometric analysis. Both temperature and pH had strongly affected the secondary structure of β-casein, with most affected regions involving random coils and α-helical structures. The α-helical structures showed great pH sensitivity by decreasing at 20 °C and diminishing completely at 4 °C when pH was increased from 5.9 to 7.0. The decrease in α-helix was likely related to the greater presence of random coils at pH 7.0, which was not observed at pH 5.9 at either temperature. The changes in secondary structure components were linked to decreased hydrophobic interactions at lower temperature and increasing pH. The most prominent change of the α-helix took place when the pH was adjusted to 7.0 and the temperature set at 4 °C, which confirms the disruption of the hydrogen bonds and weakening of hydrophobic interactions in the system. The findings can assist in establishing the structural behaviour of the β-casein under conditions that apply as important for solubility and production of β-casein.

Keywords: FTIR; NMR; β-casein; secondary structure; temperature; pH

check for updates

Citation: Markoska, T.; Daniloski, D.; Vasiljevic, T.; Huppertz, T. Structural Changes of β-Casein Induced by Temperature and pH Analysed by Nuclear Magnetic Resonance, Fourier-Transform Infrared Spectroscopy, and Chemometrics. *Molecules* **2021**, *26*, 7650. https://doi.org/10.3390/molecules26247650

Academic Editor: Mirella Nardini

Received: 4 December 2021
Accepted: 15 December 2021
Published: 17 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Caseins represent the major protein constituents of the bovine milk (80%), comprised of four major phosphoproteins, α_{S1}-, α_{S2}-, β-, and κ-casein at concentration of 12–15, 3–4, 9–11 and 2–4 g/L, respectively [1]. In the dairy industry, the structural organisation of caseins can have a great effect on the functional properties of milk and the quality of the final product due to various interactions (hydrophobic, electrostatic, van der Waals forces, or covalent bonds) yielding to different properties. As the most amphiphilic, bovine β-casein represents 35% of the caseins in the bovine milk with 209 amino acids in the polypeptide chain and average molecular weight of 24 kDa [2]. Due to the presence of many relatively hydrophobic parts in the β-casein molecule with great number of Pro residues, the molecule can adopt flexible conformations. These conformations are characterised by large range of inter or intra molecular interactions, fewer secondary and tertiary structures, while containing more random coil organisations [3].

There have been multiple studies analysing the secondary structure of β-casein. The summarised findings included the presence of 7–25% α-helical structures, 15–33% β-sheets,

Molecules **2021**, *26*, 7650

20–30% turns, and 20–25% polyproline II structures [2,4–6]. The predicted α-helical structures are likely in the N-terminal part of the β-casein, f(1–40), due to the presence of phosphoserine residues that carry the net charge of the molecule [7]. The apolar residues present in the C-terminal part of β-casein, f(136–209), are responsible for the appearance of β-sheet structures due to higher hydrophobicity in this region [7,8].

Temperature is an important factor in regards to the β-casein release from the casein micelle. In a cold environment, due to weakening of hydrophobic attraction, selective solubilisation of β-casein takes place, leading to increased solubility of β-casein during cold storage [9]. At low temperature, β-casein appears as monomer with structural transitions due to cold denaturation. The dissociation of β-casein from the casein micelle was previously observed to be the greatest at pH 5 and a temperature below 4 °C [10]. Moreover, the net negative charge at high pH increases as a result of deprotonation of carboxyl groups and loss of the positive charge of amino acids including His. The increase in electrostatic repulsion takes place as a result of the negative charge of the phosphoseryl residues [11]. The transition that takes place on the β-casein molecule based on pH and temperature adjustment can have a significant effect on the secondary structure of the molecule. The impact of both pH and temperature on β-casein and overall milk protein structure play an important role in protein stability during processing of milk and milk products and their storage. The pH of dairy products may vary from <4 for some fermented dairy to >7 for some dairy ingredients and infant formula products, whereas storage occurs both at ~4 °C for pasteurised products and at room temperature for sterilised products.

Even though several studies already confirmed the presence of some structural elements in the β-casein molecule, it would be valuable to know how this particular casein adapts its secondary structural elements in different environments [4,12,13]. Thus, the aim of the present research was to identify the structural components of β-casein using Fourier-transform infrared spectroscopy (FTIR) and nuclear magnetic resonance (NMR). FTIR spectroscopy provides information of the structural changes in proteins by absorption of the stretching/bending vibrations in different regions [14]. Consequently, NMR provides detailed information of the protons in the molecule, which evaluates the primary position of the atoms in the molecule [15]. Both methods are nondestructive with minimal sample preparation and can provide information under any condition. The current work involved structural analysis of β-casein at two pH values (pH 5.9 and pH 7.0) and two temperatures (4 and 20 °C) in order to establish the structural transitions of β-casein. The selected temperature was used to observe the structural transition of β-casein in most used storage conditions for milk proteins. The pH 7.0 was used to observe how β-casein structure changes in pH environment close to native pH in milk and pH 5.9 were used as the native pH of synthesised β-casein and also were important for curd stability during milk coagulation.

2. Results and Discussion

In the current study, FTIR and NMR were used to examine the structural changes in β-casein resulting from changes in pH and temperature. In the Amide I region of the FTIR spectra, six structural features were distinguished and analysed (Table 1).

From the peak area calculations of the second derivative spectra, the most significant differences ($p \leq 0.05$) as a result of changes in pH and temperature were observed in random coil, α-helix, and β-turn structures. At pH 7.0, the percentage of random coil structure was ~15% higher at 4 °C than at 20 °C, whereas random coil structures were not detected at pH 5.9 at either temperature. The assigned random coil structures have previously been shown to include short polyproline II (PPII) helix/chains [16]. Bovine β-casein contains 35 Pro residues, evenly distributed along the 209 amino acid polypeptide chain, which favours the formation of PPII structures [4]. The lack of detected random coil structures at pH 5.9 indicates that there was no substantial unfolding of the secondary structure of the protein at this pH [17]. The absence of random coils in β-casein at pH 5.9 appeared to correlate with a greater amount of α-helical structures, at both 4 and 20 °C. At

20 °C, the α-helix content at pH 7.0 was half of that observed at pH 5.9, whereas α-helical structures were not detected at pH 7.0 at 4 °C. The β-turn structures were affected by temperature, but not by pH; at both pH 5.9 and 7.0, fewer β-turns were observed at 20 than at 4 °C. Farrell et al. [4] stated that with the temperature increase the conversion from β-turns and β-sheet to PPII conformations might occur. Hence, in line with these previous findings in combination with the present data of random and turn conformations (Table 1), there is a reasonable possibility of transformation of these structures into PPII helices, whose content was found to be between 20 and 25% within the β-casein molecule [5,6]. The formation of β-turns in caseins is promoted by the Pro residues by cross-linking with neighbouring residues via van der Waals interactions acting as a β-turn inducer [8].

Table 1. Total percentage areas of different secondary structures in Amide I in β-CN in FTIR including side chain, β-sheet, random coil, α-helix β-turn, and aggregated β-sheets. The selected band frequency for each structural component is presented in cm^{-1}. The peak area percentage is presented for temperature of 4 and 20 °C and pH of 5.9 and 7.0.

Band Assessment	Band Frequency (cm^{-1})	Peak Area (%)			
		Temperature 4 °C		Temperature 20 °C	
		pH 5.9	pH 7.0	pH 5.9	pH 7.0
Side chain	1608–1610	3.99 ± 0.53 ^{a}	5.40 ± 0.09 ^{a}	3.92 ± 0.34 ^{a}	4.33 ± 0.60 ^{a}
β-sheet	1620–1630	26.04 ± 3.55 ^{a}	23.07 ± 4.24 ^{a}	26.70 ± 4.03 ^{a}	23.72 ± 2.47 ^{a}
Random coil	1640–1645	n/d	46.53 ± 5.81 ^{b}	n/d	31.07 ± 8.71 ^{a}
α-helix	1646–1652	42.98 ± 3.41 ^{b}	n/d	45.24 ± 3.76 ^{b}	21.77 ± 2.59 ^{a}
β-turn	1677–1679	22.25 ± 1.87 ^{c}	20.04 ± 1.35 ^{bc}	17.00 ± 2.29 ^{ab}	13.34 ± 4.66 ^{a}
Aggregated β-sheet	1689–1690	4.75 ± 1.61 ^{a}	4.96 ± 0.97 ^{a}	7.15 ± 1.36 ^{a}	5.78 ± 2.02 ^{a}

^{a,b,c} Mean values within a row that do not share a common superscript letter are significantly different ($p \leq 0.05$); n/d = not detected.

The observed differences in the peak area (Table 1) were further confirmed by PCA results (Figure 1).

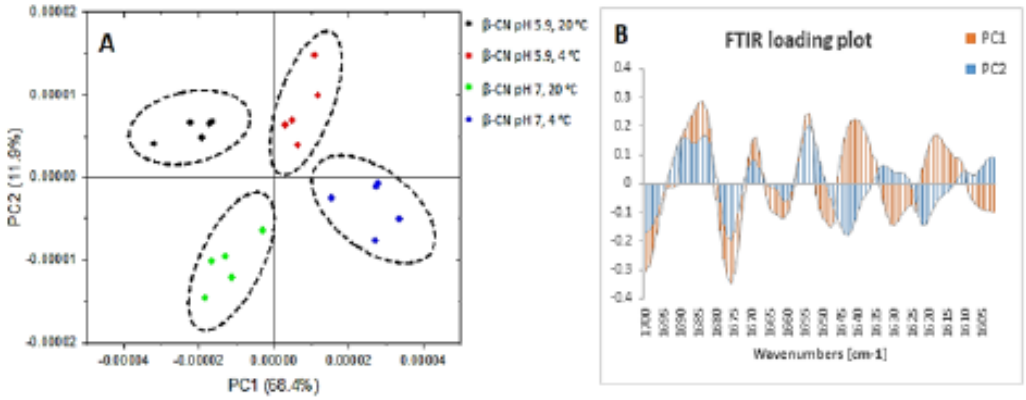


Figure 1. Principal components score plot (**A**) and loading plot (**B**) of β-casein of the FTIR spectra in region 1700–1600 cm^{-1} where β-casein at pH 5.9 and temperature of 20 °C (black), β-casein at pH 5.9 and temperature of 4 °C (red), β-casein at pH 7.0 and temperature of 20 °C (green) and β-casein at pH 7.0 and temperature of 4 °C (blue) are in (**A**) and PC1 (orange) and PC2 (blue) are in (**B**).

In addition, PC1 confirmed 68.4% difference between the samples at 4 and 20 °C, while PC2 confirmed 11.9% difference between samples at pH 5.9 and pH 7.0 (Figure 1A). The loading score in Figure 1A visually separated the groupings of samples that differed in temperature along the PC1. The PC1 also distinguished the differences in the structural components between these samples, which are confirmed in the loading plot in Figure 1B.

Molecules **2021**, *26*, 7650

The PC1 identified high loading for a peak at 1650 cm^{-1} (α-helix) for samples at 20 °C, confirming a greater presence of this secondary structure at this temperature. On the other hand, the observed high loading for peak at 1640 cm^{-1} for samples at 4 °C confirmed the high level of random structures. PC1 also depicted slight changes in β-sheets by greater presence of intramolecular β-sheets (1630 cm^{-1}) for samples at 20 °C and greater intermolecular β-sheets (1620 cm^{-1}) for samples at 4 °C. Similarly, PC2 differentiated the structural features occurring between samples with different pH values which were also present with different loadings in the loading plot in Figure 1B. In the PC2 loading plot (Figure 1B), great loadings were observed for a peak at 1653 cm^{-1} (α-helix) for samples at pH 5.9 and at 1643 cm^{-1} (random coils) for samples at pH 7.0. Regarding the β-sheet structures, greater loading for a peak at 1630 cm^{-1} (intramolecular β-sheets) was observed for samples at pH 5.9 and at pH 7.0 greater loading shifted to indicate intermolecular β-sheets (peak 1623 cm^{-1}). Thus, PC1 and PC2 successfully confirmed the previously observed changes for α-helix and random structures and additionally confirmed differences in inter- and intramolecular β-sheets.

The ^{1}H-NMR chemical shifts in proteins are sensitive to local structural rearrangements. Moreover, the conformational differences can be a result of multiple contributions including torsion angles coming from backbone and side chains, hydrogen bonding, electric fields, rings vibrations, and steric repulsions [18]. In the current work, we observed the variations in chemical shift distribution in four regions including methyl -CH_3- (0–1.5 ppm), aliphatic -CH_2- (1.5–3.5 ppm), H^{α} amide region (3.5–5 ppm), and aromatic/H^{N}-amino region (5.5–10 ppm). From the overlayed spectra in Figure 2A, a downfield shift of the proton chemical shifts can be seen at 4 °C compared to at 20 °C. In addition, the deshielding was observed in all the regions, except in the amino or backbone region (Figure 2A). Thus, in the amino region the peaks appeared in a broad pattern which overlayed in identical chemical shift positions (Figure 2A).

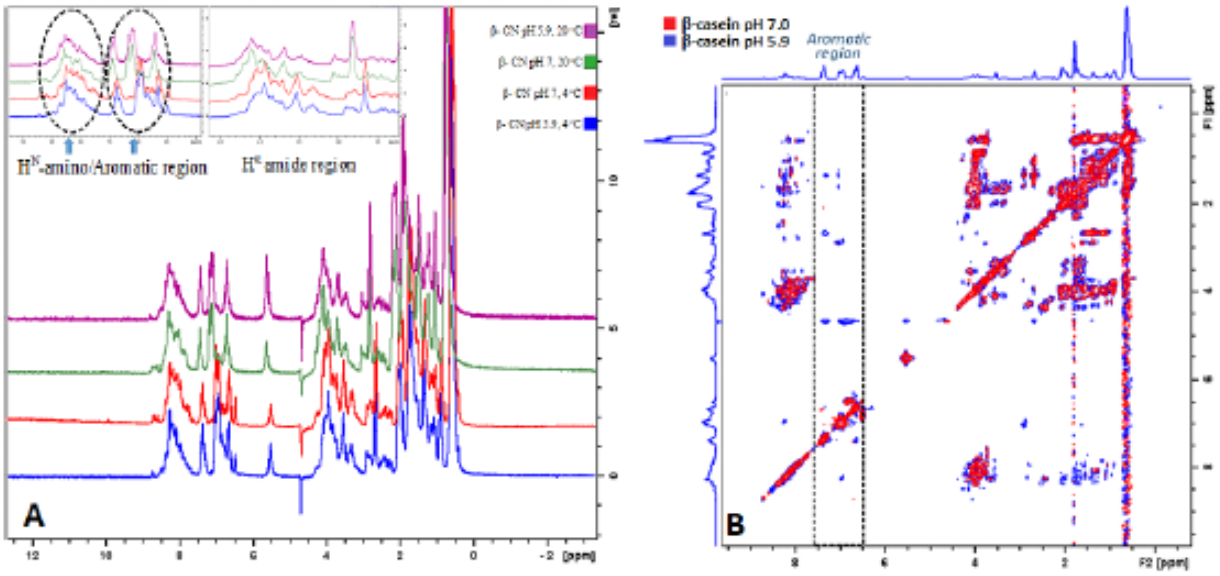


Figure 2. (**A**) Overlayed ^{1}H NMR spectra of β-casein at pH 5.9 and temperature of 20 °C (purple), β-casein at pH 5.9 and temperature of 4 °C (blue), β-casein at pH 7.0 and temperature of 20 °C (green) and β-casein at pH 7.0 and temperature of 4 °C (red). (**B**) Overlayed TOCSY NMR spectra of β-casein at pH 7.0 (red) and β-casein at pH 5.9 (blue).

The deshielding in the other regions resulted by lowering of temperature to 4 °C was observed to be 0.2–0.3 ppm. The sensitivity of the ^{1}H chemical shift can be promoted by multiple factors, including the protein sensitivity to conformations, hydrogen bonding, ring vibrations, or electric fields [19]. In addition, the upfield and downfield shielding of the ^{1}H chemical shift can result from changes in the different components in the secondary

structure [18,20]. In this study, an opposite effect by 0.3 ppm downfield shift for the amide region was observed, which can relate to a lower presence of α-helical structures at low temperature. In addition, from the integrated amide region we observed a significant decrease ($\leq$0.05) in the peak intensities by ~3% at 4 °C compared to at 20 °C (Table 2). This further confirms previously discussed FTIR results (Figure 1 and Table 1) regarding the absence or decline in α-helical structures at low temperature.

Table 2. Total percentage areas peaks in different regions for β-casein in NMR including methyl, aliphatic, amide and amino region. The difference in the percentage is presented for β-casein at temperature of 4 and 20 °C and pH of 5.9 and 7.0.

Temperature	**4 °C**		**20 °C**	
pH	**5.9**	**7.0**	**5.9**	**7.0**
Region/Integral	**β-Casein**			
Methyl	19.48 ± 0.16 ^{a}	17.63 ± 0.27 ^{b}	16.78 ± 0.13 ^{bc}	16.22 ± 0.17 ^{c}
Aliphatic	16.22 ± 0.16 ^{ab}	17.73 ± 0.75 ^{a}	14.91 ± 1.64 ^{b}	13.90 ± 0.35 ^{bc}
Amide	42.29 ± 0.04 ^{b}	42.82 ± 0.21 ^{b}	46.69 ± 0.14 ^{a}	46.32 ± 0.17 ^{a}
Amino	22.01 ± 0.34 ^{b}	21.83 ± 0.82 ^{b}	22.62 ± 0.41 ^{ab}	23.47 ± 0.89 ^{a}

^{a,b,c} Mean values within a row that do not share a common superscript letter are significantly different ($p \leq 0.05$).

From the other regions of the NMR spectrum a greater amount of methyl groups at low temperature was observed, which was more present at pH 5.9 (Figure 2A). The peak observed in this region (0–0.8 ppm) likely appeared from the side chains (CH_3) groups from Ile, Val, and Leu in β-casein [21]. The lower peak intensity for methyl groups at room temperature can be due losses of signal from the methyl side chains of Ile, Val, and Leu. An increase in pH from 5.9 to 7.0 led to slightly lower peaks for the methyl side chains (Table 2). The amide region of the β-casein presented slightly greater intensity at 20 °C compared to that at 4 °C at both pH values (Table 2).

The change in pH from 5.9 to 7.0 lead to lower detection of the cross peaks in the aromatic region of the TOCSY spectra (Figure 2B). This spectral pattern was not affected by the temperature, i.e., at both 4 and 20 °C the TOCSY spectra of β-casein showed the same pattern in regard to the pH. In the aromatic region of the TOCSY spectra the cross peaks present the interactions of the proton rings with the neighbouring protons of the amino acids [15]. It was previously suggested that charge-induced disruption of the cohesive interactions in the hydrophobic regions of the caseins occur as a result of pH-induced charge modifications on side groups of amino acids, including Tyr, Lys, His, and Arg [22]. In the pH range currently studied, His would be the most susceptible to (de)protonation as a result of pH change. The side groups of these amino acids are detected in the aromatic region of the TOCSY spectra and the cross peaks in this region denote their connection with the neighbouring protons. It is known that increase in pH leads to greater intermolecular repulsion of caseins [11]. The addition of alkali to the β-casein solution can lead to disruption of existing hydrophobic connections, which was observed by reduction of the proton–proton cross association in the TOCSY spectra.

The observed structural differences resulted from changes in temperature and pH were also evaluated by PCA analysis (Figure 3).

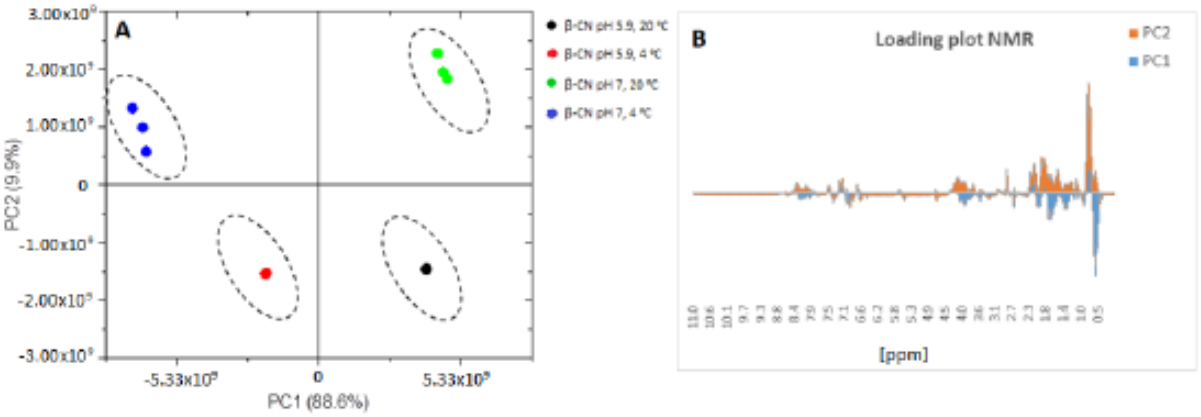


Figure 3. Principal components score plot (**A**) and loading plot (**B**) of β-casein of the NMR spectra in where β-casein at pH 5.9 and temperature of 20 °C (black), β-casein at pH 5.9 and temperature of 4 °C (red), β-casein at pH 7.0 and temperature of 20 °C (green) and β-casein at pH 7.0 and temperature of 4 °C (blue) are in (**A**) and PC1 (blue) and PC2 (orange) are in (**B**).

In the PCA loading score (Figure 3A), PC1 separated the results based on the temperature with 88.6% variance. On the other hand, PC2 separated the β-casein based on pH effect with 9.9% variance. PC1 has shown greater loading for a peak at 0.5 ppm for samples at low temperature and PC2 indicated that this peak was mainly present for samples at pH 7.0. Moreover, PC2 showed that peak loadings in aliphatic (1.5–3.5 ppm) and amide regions (3.5–5.0 ppm) were more prominent for samples at pH 7.0, whereas at pH 5.9 greater loading was observed for the aromatic region (6.5–8.0 ppm). Furthermore, both PC1 and PC2 confirmed that in the aliphatic and aromatic regions the peak loading was more intense for samples at pH 7.0 and temperature of 4 °C, which can confirm the structural changes observed by FTIR occurring as a result of a cold environment and pH.

3. Material and Methods

3.1. Sample Preparation

Bovine β-casein was purchased from Sigma-Aldrich (Sigma-Aldrich, St. Louis, MO, USA) with >98% purity. The preparation process was performed by dispersion of 10 mg β-casein in 1 mL of a mixture of H_2O:D_2O (90%:10%, Sigma-Aldrich, St. Louis, MO, USA) for both FTIR and NMR analyses. The pH of β-casein in H_2O/D_2O solution was measured using a pH meter (Metrohm AG, Herisau, Switzerland) equipped with a combined pH electrode with temperature sensor and fixed cable. The original pH of the protein solution was 5.9. The pH was further adjusted to 7.0 using 0.25 M NaOH. The pH adjustment was carried out under continuous stirring of the solution and pH was additionally controlled before the structural measurements. Both β-casein solutions with pH 5.9 and pH 7.0 were analysed for structural features at 4 and 20°C. The temperature for the FTIR analysis was monitored with temperature probe, however for the NMR analysis the temperature control was automated by the instrument.

3.2. Fourier Transform Infrared Spectroscopy

The secondary structure of β-casein as a function of pH and temperature was analysed using a FTIR spectrometer (Frontier, PerkinElmer, Boston, MA, USA) in the range of 4000 to 600 cm^{-1} with a resolution of 4 cm^{-1} and 16 scans for each spectrum. Before the start of the measurement for every sample, the background spectra were scanned using a blank diamond attenuated total reflectance (ATR) cell. At the start of the analysis the solvent spectrum (90% H_2O/10% D_2O) was recorded and used for subtraction from the sample spectra to eliminate the intense solvent signal. The spectra of five subsamples of each sample were taken by refilling the ATR cell. For the secondary structure analysis, the amide I region (1700–1600 cm^{-1}) was analysed for C=O stretching after

applying Savitzky–Golay smoothing and 2nd derivative. The FTIR spectra and mean centering were analysed using Spectragryph software (version 1.2.7, Oberstdorf, Germany). The baseline of five repetitions for each sample was subtracted and the area of the peak was analysed by Origin software (Origin Pro 2021, v. 95E, OriginLab Corporation, Northampton, MA, USA). To estimate the area of each component representing secondary structures, a technique previously described elsewhere [23] was utilised. Briefly, each spectrum from five repetitions was treated by interpolation baseline mode and adjacent-averaging smoothing, then each peak in the 2nd derivative spectra was selected and integrated by high percentage. In the amide I region several peaks were analysed, including side chain (1608–1610 cm^{-1}), β-sheets (1620–1630 cm^{-1}), random coil (1640–1645 cm^{-1}), α-helix (1646–1652 cm^{-1}), β-turns (1677–1679 cm^{-1}) and aggregated β-sheets (1689–1690 cm^{-1}) [24].

3.3. Nuclear Magnetic Resonance

The NMR analysis was performed on a Bruker Avance spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany) operating at a 600 MHz transmitter frequency using 5 mm TXI probe with z- gradient. Proton NMR (^{1}H-NMR) spectra were acquired using 16 scans and spectral width of 9615 Hz in three replications. The two-dimensional methods used in this study were total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY). The 2D spectra were recorded using 8 scans and spectral width of 8196 Hz for TOCSY and 5882 Hz for NOESY spectra. The water suppression for spectra was performed using excitation sculpting with gradients with acquisition mode of States-TPPI (time-proportional phase incrementation). All spectral data were processed using TopSpin (version 4.0.6) software (Bruker BioSpin). The phase correction was performed manually by either 0th or 1st order correction for pk or the baseline and the noise correction was adjusted using qfil mode to 0.1 ppm filter width and symmetrical noise correction for the homonuclear spectra. The NMR spectra were analysed using three NMR regions, i.e., amino/aromatic (H^N), amide (H^α), and aliphatic. From the aliphatic region, the methyl peak (CH_3) was analysed separately due to its high loading. All the peaks observed in these regions were manually integrated in TopSpin (version 4.0.6) software (Bruker BioSpin).

3.4. Statistical Analysis and Spectral Data

To assess the experimental results, Minitab version 20 software was used to analyse the data (Minitab Inc., State College, PA, USA). Hence, a two-way ANOVA and Tukey's test were performed with β-casein as fixed factor to evaluate if there was a difference between means ($p \leq 0.05$). Unless otherwise specified, all measurements were taken in triplicate. In addition, principal component analysis (PCA) was used for both FTIR and NMR results to identify the changes in the secondary structure in β-casein samples based on the temperature and pH effect. PCA gives information by generating principal components (PCs) as a coordinated axis with a least possible loss of information. The groupings of the different samples were depicted in score plots and the loading plots were used for the wavenumbers responsible for peaks classification. The multivariate analysis was carried out at a 95% confidence level.

4. Conclusions

The combination of FTIR and NMR was very efficient in characterisation of structural differences in the β-casein molecule as affected by temperature and pH. Moreover, both methods confirmed similar structural transitions of the β-casein molecule in regard to the applied conditions, indicating greater structural changes taking place as a result of the temperature. A temperature of 4 °C is critical for β-casein molecules due to lowering of hydrophobicity that results in unfolding and opening up of the native conformers and liberalisation of the monomers from the casein micelle. The restructuring of the β-casein was confirmed by FTIR results implying formation of more random coils at 4 °C at the expense of the α-helical structure. Moreover, this was intensified by changing the pH

to 7.0. Hence, the side chains of amino acids known as α-helix inducers showed lower detection due to loss of hydrogen bonding resulting in reduction of the α-helix in the β-casein polypeptide chain. The observed structural changes promoted by the temperature change were more intense at high pH, however at pH 5.9 the structural reorganisation was minor. The current findings confirmed that both pH and temperature have a great effect on structure of a β-casein molecule. Moreover, the results can assist in understanding the behavior of β-casein during β-casein production, solubility, processing of milk and milk products, and storage. In addition, FTIR and NMR have proven to successfully detect structural transition in β-casein. The combination of these methods can be further expanded to other proteins and assist in understanding protein behavior during food processing.

Author Contributions: T.M. conceived the study and research question; designed and wrote the original draft, conceptualised, reviewed, and edited the manuscript. D.D. gave critical feedback and analysis, and reviewed and edited the manuscript. T.H. and T.V. provided critical feedback and analysis, secured funding, reviewed and edited the manuscript, and supervised the study. All authors have read and agreed to the published version of the manuscript.

Funding: The first author is funded by a Victoria University Research Training Stipend Scholarship (Grant number—INT010).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

Sample Availability: Samples of the compounds are not available from the authors.

References

1. Huppertz, T. Chemistry of the caseins. In *Advanced Dairy Chemistry*; Springer: Boston, MA, USA, 2013; pp. 135–160. [CrossRef]
2. Huppertz, T.; Fox, P.; Kelly, A. The caseins: Structure, stability, and functionality. In *Proteins in Food Processing*, 2nd ed.; Yada, R.Y., Ed.; Woodhead Publishing: Sawston, UK, 2018; pp. 49–92. [CrossRef]
3. Zhou, Z.; Zhu, M.; Zhang, G.; Hu, X.; Pan, J. Novel insights into the interaction mechanism of 5-hydroxymethyl-2-furaldehyde with β-casein and its effects on the structure and function of β-casein. *LWT* **2021**, *152*, 112360. [CrossRef]
4. Farrell, H.; Qi, P.; Wickham, E.; Unruh, J. Secondary structural studies of bovine caseins: Structure and temperature dependence of β-casein phosphopeptide (1-25) as analyzed by circular dichroism, FTIR spectroscopy, and analytical ultracentrifugation. *J. Protein. Chem.* **2002**, *21*, 307–321. [CrossRef]
5. Qi, P.X.; Wickham, E.D.; Farrell, H.M. Thermal and Alkaline Denaturation of Bovine β-Casein. *Protein. J.* **2004**, *23*, 389–402. [CrossRef] [PubMed]
6. Qi, P.X.; Wickham, E.D.; Piotrowski, E.G.; Fagerquist, C.K.; Farrell, H.M. Implication of C-terminal deletion on the structure and stability of bovine β-casein. *Protein. J.* **2005**, *24*, 431–444. [CrossRef] [PubMed]
7. Creamer, L.; Richardson, T.; Parry, D. Secondary structure of bovine αs1-and β-casein in solution. *Arch. Biochem. Biophys.* **1981**, *211*, 689–696. [CrossRef]
8. Kumosinski, T.; Brown, E.; Farrell, H., Jr. Three-dimensional molecular modeling of bovine caseins: An energy-minimized β-casein structure. *J. Dairy Sci.* **1993**, *76*, 931–945. [CrossRef]
9. Huppertz, T.; Hennebel, J.-B.; Considine, T.; Kelly, A.L.; Fox, P.F. A method for the large-scale isolation of β-casein. *Food Chem.* **2006**, *99*, 45–50. [CrossRef]
10. Post, A.; Arnold, B.; Weiss, J.; Hinrichs, J. Effect of temperature and pH on the solubility of caseins: Environmental influences on the dissociation of αS-and β-casein. *J. Dairy Sci.* **2012**, *95*, 1603–1616. [CrossRef]
11. Horne, D.S. Casein interactions: Casting light on the black boxes, the structure in dairy products. *Int. Dairy J.* **1998**, *8*, 171–177. [CrossRef]
12. Condict, L.; Kaur, J.; Hung, A.; Ashton, J.; Kasapis, S. Combined spectroscopic, molecular docking and quantum mechanics study of β-casein and ferulic acid interactions following UHT-like treatment. *Food Hydrocoll.* **2019**, *89*, 351–359. [CrossRef]
13. Farrell Jr, H.; Wickham, E.; Unruh, J.; Qi, P.; Hoagland, P. Secondary structural studies of bovine caseins: Temperature dependence of β-casein structure as analyzed by circular dichroism and FTIR spectroscopy and correlation with micellization. *Food Hydrocoll.* **2001**, *15*, 341–354. [CrossRef]
14. Gallagher, W. FTIR analysis of protein structure. *Biochemistry* **1997**, *392*, 662–666.

15. Markoska, T.; Vasiljevic, T.; Huppertz, T. Unravelling Conformational Aspects of Milk Protein Structure—Contributions from Nuclear Magnetic Resonance Studies. *Foods* **2020**, *9*, 1128. [CrossRef] [PubMed]
16. Dukor, R.K.; Keiderling, T.A. Reassessment of the random coil conformation: Vibrational CD study of proline oligopeptides and related polypeptides. *Biopolymers* **1991**, *31*, 1747–1761. [CrossRef]
17. Nishinari, K.; Zhang, H.; Ikeda, S. Hydrocolloid gels of polysaccharides and proteins. *Curr. Opin. Colloid. Interface. Sci.* **2000**, *5*, 195–201. [CrossRef]
18. Shen, Y.; Bax, A. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR* **2007**, *38*, 289–302. [CrossRef]
19. Mulder, F.A.; Filatov, M. NMR chemical shift data and ab initio shielding calculations: Emerging tools for protein structure determination. *Chem. Soc. Rev.* **2010**, *39*, 578–590. [CrossRef]
20. Wang, Y.; Jardetzky, O. Investigation of the neighboring residue effects on protein chemical shifts. *J. Am. Chem. Soc.* **2002**, *124*, 14075–14084. [CrossRef]
21. De Angelis Curtis, S.; Curini, R.; Delfini, M.; Brosio, E.; D'Ascenzo, F.; Bocca, B. Amino acid profile in the ripening of Grana Padano cheese: A NMR study. *Food Chem.* **2000**, *71*, 495–502. [CrossRef]
22. Vaia, B.; Smiddy, M.A.; Kelly, A.L.; Huppertz, T. Solvent-mediated disruption of bovine casein micelles at alkaline pH. *J. Agric. Food. Chem.* **2006**, *54*, 8288–8293. [CrossRef]
23. Daniloski, D.; McCarthy, N.A.; Markoska, T.; Auldist, M.J.; Vasiljevic, T. Conformational and physicochemical characteristics of bovine skim milk obtained from cows with different genetic variants of β-casein. *Food Hydrocoll.* **2022**, *124*, 107186. [CrossRef]
24. Markoska, T.; Huppertz, T.; Grewal, M.K.; Vasiljevic, T. FTIR analysis of physiochemical changes in raw skim milk upon concentration. *LWT* **2019**, *102*, 64–70. [CrossRef]

CHAPTER 6: Changes in the secondary structure of β-casein from A1 and A2 milk in regards to the pH and genetic alternative

Abstract

The study investigated the structural differences between β-casein (CN) from A1 and A2 milk in pH conditions as in raw milk (pH 6.7) and at isoelectric point of β-CN (pH 4.6). The aim of this study is to detect the main reason for the differences in the coagulation properties between A1 and A2 milk for which the main distinguishing factor is the different amino acid in the β-CN molecule. For the purpose, spectroscopic tools were coupled with chemometric analysis. This included Nuclear Magnetic Resonance (NMR), Fourier Transform Infrared (FTIR) and Raman spectroscopy. In addition, the physical appearance of the β-CN molecule was observed by analysis of the particle size and microscopic imaging. The results in this work shown that β-CN from A2 milk have open and porous structure and at pH of 4.6 forms irregular aggregations with predominate random coils or polyproline II structures. The β-CN from A1 milk have more dense protein network compared to β-CN from A2 milk. At pH 4.6 the aggregations are composed of aggregated β-sheets components and small network cavities. The additional proline in β-CN from A1 milk affects the structure of the casein and thus the functional properties of the milk. The findings in this work will assist the industry and researchers in understanding the functional properties of A1 and A2 milk based on the structural arrangements of β-CN from these two milk genotypes.

1. Introduction

Caseins are milk proteins that have important role as a nutritive component for neonate grow and play crucial role in gelation process during production of fermented dairy products (Lucey, 2002). There are four caseins in the milk including αs_1-, αs_2-, β-, and κ-casein (CNs) which form a complex micellar structure held together by colloidal calcium phosphate nanoclusters. β-CN represent ~40 % of the total caseins in milk (Huppertz, 2012). This casein have important role in the acid and rennet gelation of the milk due to the chaperone activity in milk and ability to act as surfactant (Nguyen *et al*., 2015: Zhang *et al*., 2018; Daniloski *et al*., 2022). The impact of β-CN on milk gelation properties is dependent on the genetic variants. Moreover, there have been observed fifteen genetic variants of β-CN from which A1 (β-CN-A1) and A2 (β-CN-A2) are the most common and known to the public. The impact of the A1 and A2 β-CN variant on the milk functionality and gelation properties have raised great interest among the industries, researchers and consumers (Poulsen *et al*., 2013; Daniloski *et al*., 2022). The primary difference between β-CN-A1 and β-CN-A2 is in the position 67 of the polypeptide chain. Thus, in β-CN-A1 this position is occupied by histidine and in β-CN-A2 in position 67 is present proline (Daniloski *et al*., 2021).

The difference in one amino acid in the polypeptide chain can have great impact on structural transition of the proteins during milk processing. In our previous study we observed the structure of peptide with 11 amino acids (β-casomorphin 11 (BCM11)) which contained identical amino acid sequence as in BCM11-A1 and BCM11-A2 (Markoska *et al*., 2023). The findings shown that BCM11-A1 have less dense structure with exposed hydrophilic parts making available sites for hydrogen bond formation. However, BCM11-A2 was observed to have more packed and irregular structure. The dissimilarities in the structure of the peptides can also reflect on structural variations of β-CN molecule and thus its functionality in milk processing. The yoghurt produced from A1 and A2 β-CN variant have different protein networks. Thus, yoghurt from β-CN-A2 results with more porous gel with lower strength compared to β-CN-A1 yoghurt gel (Nguyen *et al*., 2018). The gelation properties of A1 and A2 milk using acid-induced gelation confirmed more porous gels and smaller particles for A2 milk compared to A1 milk (Daniloski *et al*., 2022), The smaller particles for A2 variant were observed to be

due to chaperone activity of β-CN-A2 promoted by monomer transition (Raynes *et al*., 2015). This behaviour of β-CN-A2 was postulated to be a result of more polyproline (PPII) helixes in the molecule compared to β-CN-A1 (Raynes *et al*., 2015). However, there have not been direct evidence of the secondary structure of β-CN from A1 and A2 milk under conditions as found in milk or when aggregated. Therefore, in this work we observed the transition of the secondary structure of β-CN-A1 and β-CN-A2 at pH 6.7 (native pH of milk) and pH 4.6 (isoelectric point).

The study was conducted using multiple spectroscopic tools including Nuclear Magnetic Resonance (NMR), Fourier Transform Infrared (FTIR) spectroscopy and Raman spectroscopy all combined with chemometrics in order to identify the structural differences between β-CN-A1 and β-CN-A2. In addition, the particle size and microstructure of the casein particles were observed to compliment the findings of the spectroscopic tools.

2. Materials and methods

2.1 Sample preparation

Raw milk was supplied by The Agriculture Victoria's Ellinbank Centre in Victoria, Australia. The cows were previously genotyped using capillary electrophoresis method as previously described by Raynes *et al*., (2015). The milk with A1 and A2 β-casein genotypes was further used for the β-CN extraction. The raw milk was skimmed by centrifugation at 3225 g for 20 min and 20 °C. β-CN was purified from the skim milk samples using the method of Huppertz *et al*., (2006) with some adjustments. Briefly, skimmed milk was warmed at 30 °C, then rennet was added and milk was held at the warming temperature for minimum of 30 minutes. The coagulum was crushed and centrifuged at 5000 g for 15 minutes at 30 °C. The whey was decanted off and replaced by hot (50-100 °C) ultra-pure water. After 5-10 minutes, the samples were centrifuge again at 5000 g for 15 minutes and 5 °C. The supernatant was discarded and the curd was macerated using mortar and pestle and then resuspended in cold (5 °C) ultra-pure water and held at 5 °C for 72 hours. After 72 hours, the suspension was mixed and centrifuged

at 5000 g for 15 minutes and 5 °C after what the supernatant was filtered and lyophilised using a freeze-drier (model FD-300, Airvac Engineering Pty. Ltd., Dandenong, Australia). The β-casein composition was confirmed to be >80 percent by reversed phase-high performance liquid chromatography (RP-HPLC) using Shimadzu system (LC-2030C - Shimadzu, Europa GmbH, Duisburg, Germany). The pH of the samples was recorded and adjusted to 4.5 and 6.5 using 0.25 M HCl or 0.25 M NaOH.

2.2 Identification of the A1 and A2 β-casein genotypes

The A1 and A2 β-casein genotypes were identified applying a reversed phase-high performance liquid chromatography (RP-HPLC) using Shimadzu system (LC-2030C - Shimadzu, Europa GmbH, Duisburg, Germany) as previously described by Daniloski *et al*., (2022). The conditions used were 45 °C and UV detector at 240 nm. The samples were prepared by resuspending the lyophilised samples in ultra-pure water from which 1.2 mL were diluted with 3.2 mL urea (Mediwaththe, Chandrapala and Vasiljevic, 2018). The mixture was then filtered using 0.45 μm syringe filters into HPLC vials. The β-CN genotypes were identified using Aeris WIDEPORE C4 column by Phenomenex (150 mm × 4.6 mm, 3.6 μm particle size, 300 Å porosity, Torrance, USA). The system used two mobile phases. Hence, mobile phase A contained 0.1% trifluoroacetic acid (TFA) in ultra-pure water and mobile phase B contained 0.1% trifluoroacetic acid (TFA) in acetonitrile (ACN). Both TFA and ACN were supplied by Sigma–Aldrich (St. Louis, MO, USA). The identification of the retention time was performed using commercially available β-CN standard (Sigma–Aldrich, St. Louis, MO, USA). The retention time was recorded at 25.64 min and 26.84 min for β-CN-A1 and β-CN-A2, respectively as reported in the previous literature (Daniloski *et al*., 2022). The identified genetic variants using HPLC are present in Supplementary Figure 1.

2.3 Nuclear Magnetic Resonance (NMR*)*

The structure of the β-CN-A1 and β-CN-A2 was further analysed using NMR by employing Bruker Avance spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany). The instrument operates at 600 MHz and for the analysis; TXI probe with z-gradient was used. For proton NMR, 50 mg of each sample was diluted in 9:1 H_2O:D_2O solution and the D2O was used to lock the signal. For ^{13}C NMR 500 mg of each sample was diluted in 9:1 H_2O:D_2O solution. The samples were analysed using 1D ^{1}H and ^{13}C.

The ^{1}H proton spectra was recorded using 16 scans and 5 replications in the spectra width of 9615 Hz. The ^{13}C spectra was recorded using 256 scans and spectra width of 32679 Hz.

The processing of the spectra was performed using a TopSpin (version 4.0.6) software (Bruker BioSpin). Thus, the baseline was corrected using qfill fid baseline mode, the FIDs were corrected by line-broadening parameter (0.3 Hz) and the phase correction was performed by 0^{th} and 1^{st} order correction for pk. The difference ^{13}C NMR spectra between spectra intensity was calculated by summing the intensity variation between the observed spectra using TopSpin (version 4.0.6) software (Bruker BioSpin). The assignments of the chemical shifts were achieved following the previous values in the literature and BMRB files (Markoska *et al*., 2021a; Seavey, Farr, Westler & Markley, 1991).

2.4 Fourier Transform Infrared (FTIR) spectroscopy

The FTIR spectra was recorded using PerkinElmer Frontier FTIR spectrometer (Frontier 1, PerkinElmer, Boston, MA, USA). From each sample around 100 µL were dropped on a diamond attenuated total reflectance (ATR) cell. The background spectrum was recorded before the start of the scanning of every sample. The spectra were recorded in a spectral range of 4000-600 cm^{-1} using 16 scans and spectral resolution of 4 cm^{-1} and for each sample 5 replications were recorded. The spectra was processed using a Spectragryph software (version 1.2.15, Oberstdorf, Germany). Solvent spectrum was recorded using ultra-pure water and was used for spectra subtraction in order to remove the intense solvent signal (H_2O) in the samples. For identification of the structural components, Amide I and Amide II regions were used in the spectral range of 1600-1700 cm^{-1} and 1600-1500 cm^{-1}, respectively. Amide I represents the C=O stretching vibrations of the peptide bond (Carbonaro and Nucara, 2010). Amide II is used to identify N-H bending and C-N stretching vibrations in the peptide (Curley *et al*., 1998). Each spectrum was processed by smoothing using Savitsky-Golay smoothing in polynomial order of 3 and 10 intervals and derivatization of each spectra was performed using 2^{nd} order of derivative.

The peaks of the absorbance spectra were analysed in Origin Pro 2021 software (v. 95 E, OriginLab Corporation, Northampton, MA, USA). For the Amide I and II Savitzky–

Golay smoothing was used and the baseline of averaged spectra from 10 individual replications was subtracted. The variability of the spectra was corrected before averaging each spectrum by calculated standard normal variates using the above-mentioned software. The averaged absorbance spectra was analysed using nonlinear curve fit (Gaussian distribution) to identify the peaks distribution in the Amide I and II region. The curve fitting had regression of 0.995, 0.998, 0.999 and 0.997 for β-CN-A1 at pH 6.7, β-CN-A1 at pH 4.6, β-CN-A2 at pH 6.7 and β-CN-A2 at pH 4.6, respectively.

The differences in the structural components in regards to the genotype and pH were further analysed using principal component analysis (PCA).The recorded spectra was evaluated in Origin Pro 2021 software (v. 95 E, OriginLab Corporation, Northampton, MA, USA). The principal components (PCs) generates information based on the individual variance including pH and genotype by providing loading scores and loading plot. The loading scores sort samples based on the variance into PC1 axis and PC2 axis, however, the loading plot providing information of the structural differences based on the groups in loading score. The PCA analysis was performed with 95 % confidence ellipse.

2.5 Raman spectroscopy

The caseins were analysed using an Alpha300 R confocal Raman microscope (WITec, Ulm, Germany), equipped with ultrafast Raman imaging SSD camera and 532 nm laser. The spectrum was collected with a laser with power of 40 mW in 20 accumulations and 1s integration time. For each sample 5 spectra were collected. The processing of the raw spectra was performed using the correction function with cosmic ray removal (dynamic factor of 8 and filter size of 3) and for the background substraction was used the shape function (noise factor of 1 and shape size of 250).

The collected spectra were analysed using Origin Pro 2021 software (v. 95 E, OriginLab Corporation, Northampton, MA, USA). Five replications of each sample were smoothed using Savitzky–Golay smoothing and the baseline was subtracted. The variability of the spectra were adjusted by calculated standard normal variates. The averaged absorbance spectra of Amide I region (1600-1700 cm^{-1}) was analysed for peak distribution using nonlinear curve fit (Gaussian distribution). The curve fitting resulted with regression of 0.998-0.999 for all observed samples.

The possible differences in the structural components in regards to the pH or genotype were analysed using principal component analysis (PCA), performed using Origin Pro 2021 software (v. 95 E, OriginLab Corporation, Northampton, MA, USA). The principal components PC1 and PC2 provides information of the sample differences in regards to the structure which results to individual groupings for samples or score plots along PC1 or PC2. The PCA analysis was performed with 95 % confidence ellipse.

2.6 Particle size

The size distribution of the samples was recorded using Zetasizer-Nano ZS (Malvern Instruments, Malvern, UK). Samples were diluted in ultra-pure water in ratio of 1:100. The refractive index was set up at 1.342 and 1.330 for β-casein and water, respectively. The data was processed using a Dispersion Technology software (version 5, Malvern Instruments). The particle size measurement was performed in five replications and every measurement was comparable to auto-correlation function recorded over period of 40s.

2.7 Scanning electron microscopy (SEM)

The microstructure of the samples was analysed using Scanning Electron Microscope (SEM) as described by Sah *et al.* (2016). Briefly, samples were snap-freeze in liquid nitrogen and kept overnight at –80 °C and freeze dried. The freeze dried samples were attached to aluminium SEM stub containing double-sided adhesive carbon tape and attached on the electron microscope stubs. The samples were then coated with chromium (K550X, Emitech, Ashford, UK). The collection of images was performed using Zeiss Supra 40P field emission SEM (Carl Zeiss SMT Ltd., Cambridge, UK) at 2.00 kV. At least 10 micrographs, collected at different random locations, were observed for each sample. The magnification of 200x was used for collecting of the representative micrographs. For each sample three replications were performed.

3. Results and discussion

3.1. Proton NMR (^{1}H)

The ^{1}H NMR spectra of the β-CN (A1 and A2) was used to observe specified regions in order to confirm the major deviation in the chemical shift distributions. The region from 3.4 to 3.9 was excluded from the analysis due to interference of the lactose signal. The βH and γH of pyrrolidine ring of proline appear at approximately 1.5-3.0 ppm (Illa *et al.*, 2017; BRMB.io). The chemical shifts in this region in β-CN-A1 casein had minor differences in regards to pH change with downfield shift of the βH chemical shift (2.5-2.8) when pH was changed from 6.7 to 4.6 (Figure 1 C,D). This deshielding was also observed for the β-CN-A2, however, change of pH from 6.7 to 4.6 promoted intense downfield shifting of the chemical shifts in the region 1.8-2.4 ppm with rearrangement and loss of chemical shifts which were primary observed at pH 6.7 (Figure 1 A,B). Moreover, in β-CN-A2 the presence of additional proline in the polypeptide chain can play a crucial role in different structural behavior as a result of pH in comparison to β-CN-A1. The chemical shift of the βH and γH regions shows greater sensitivity resulted by conformational transformations of the proline rich polypeptide chain (Ivanova *et al.*, 2010).

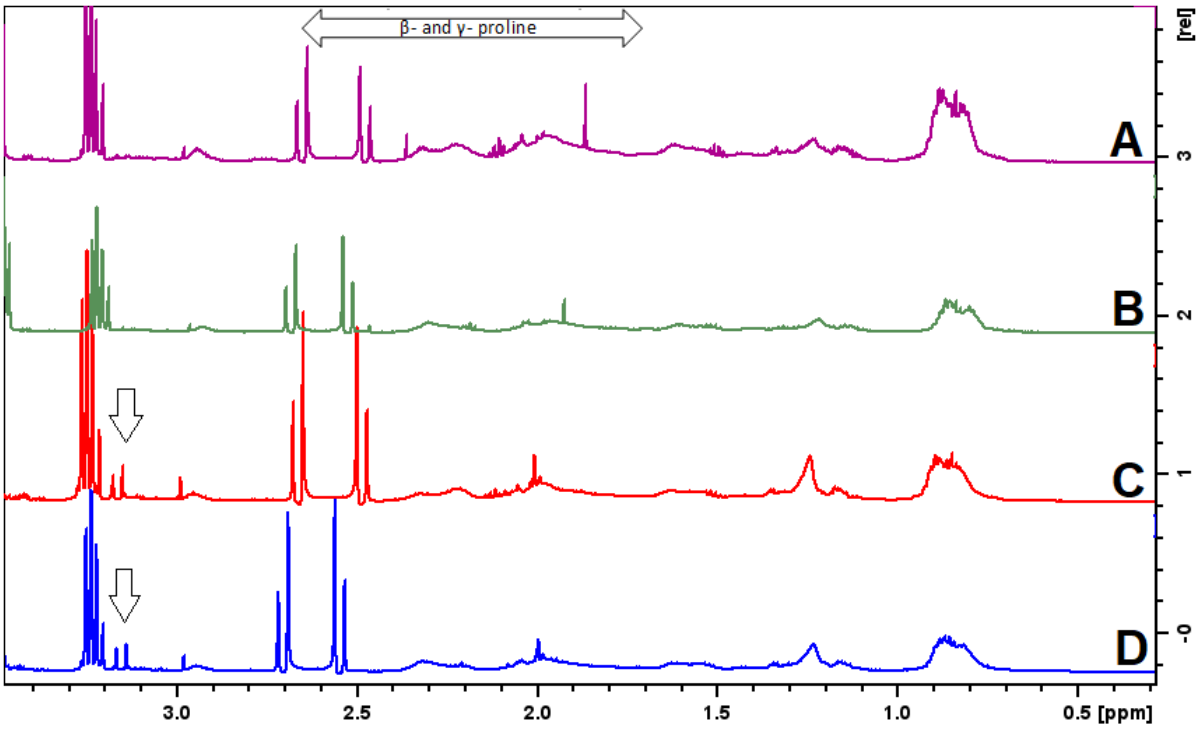


Figure 1. ^{1}H NMR spectra (0-4 ppm) of β-CN-A2 at pH 6.7 (A), β-CN-A2 at pH 4.6 (B), β-CN-A1 at pH 6.7 (C) and β-CN-A1 at pH 4.6 (D).

The different genetic variants were successfully distinguished with doublet appearing at approximately 3.15 ppm, which was assigned to βH of histidine (BRMB.io). This chemical shift appears only in β-CN-A1 generating unique distinguishing pattern for detection of the genetic variant (Figure 1 C,D). The aromatic and amino regions in the ^{1}H NMR spectra had crowded/overlapping chemical shift from multiple amino acids present in the β-CN molecule and thus is not discussed in this work.

To further discriminate the differences between the samples, the PCA analysis was used for ^{1}H NMR spectra (Figure 2). The loading score in Figure 2A represents the variance of 61.0 % for PC1 and 27.7 % for PC2. PC1 separates β-CN-A1, while PC2 separates β-CN-A2, based on two pH levels. Moreover, the loading scores shown structural differences present in the β-CNs governed by the pH value. The PCA loading plot did not show any pattern of the loadings for the observed samples owing to the size of the spectra and multiple overlapped proton chemical shifts from the 209 amino acids in the β-CN molecule (Figure 2B). However, the PCA loading score revealed different grouping of the β-CN samples due to variation of pH.

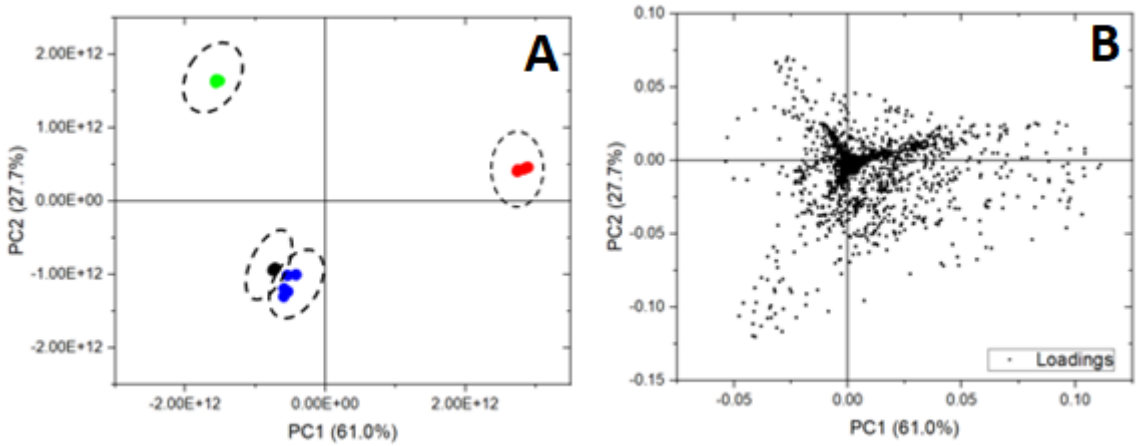


Figure 2. PCA loading score (left) for β-CN-A1 pH 6.7 (red), β-CN-A1 pH 4.6 (black), β-CN-A2 pH 6.7 (blue) and β-CN-A2 pH 4.6 (green).The loading plot (right).

3.2 ^{13}C NMR

The chemical shifts of the ^{13}C NMR signal have much weaker magnetic momentum than the proton signal, owing to inherently much weaker ^{13}C protons. The lower natural abundance of ^{13}C protons often results to difficulties in providing reasonable resonance frequency. The major chemical shifts observed in the ^{13}C NMR spectra were of the protein backbone (50-80 ppm) (Silverstein and Bassler, 1962). The aromatic region (>100 ppm) has shown only one chemical shift at 103 ppm (Figure 3). The ^{13}C NMR spectra of both β-caseins (A1 and A2) at pH 4.6 and 6.7 are shown in Figure 3. The major difference in the spectra was observed in the chemical shift intensity resulted from pH alteration. Thus, for both β-CN samples (A1 and A2) when pH was lowered from 6.7 to 4.6 the chemical shift intensity was essentially reduced. It is generally known that the intensity of the chemical shifts is directly related to the population of the species (Kleckner *et al*., 2011). Moreover, the decrease of the chemical shift intensity arises from changes in the local dynamics and structural changes (Dong *et al*., 2017). The difference between the spectra at pH 6.7 and pH 4.7 was calculated and new spectra difference was generated for both β-CN (A1 and A2) (Supplementary Figure 2). The integrated chemical shift intensity shown that the carbons from β-CN-A1 are more affected of the pH drop than the carbons from β-CN-A2. Moreover, the chemical shift intensity of the calculated spectra difference resulted from pH effect was 20-50 % higher in β-CN-A1. At pH 4.6 the β-CN molecule is electrically neutral and insoluble leading to aggregation (Huppertz, 2012). This confirms that as a result of pH changes β-CN-A1 promotes more intense structural dynamics in comparison to β-CN-A2.

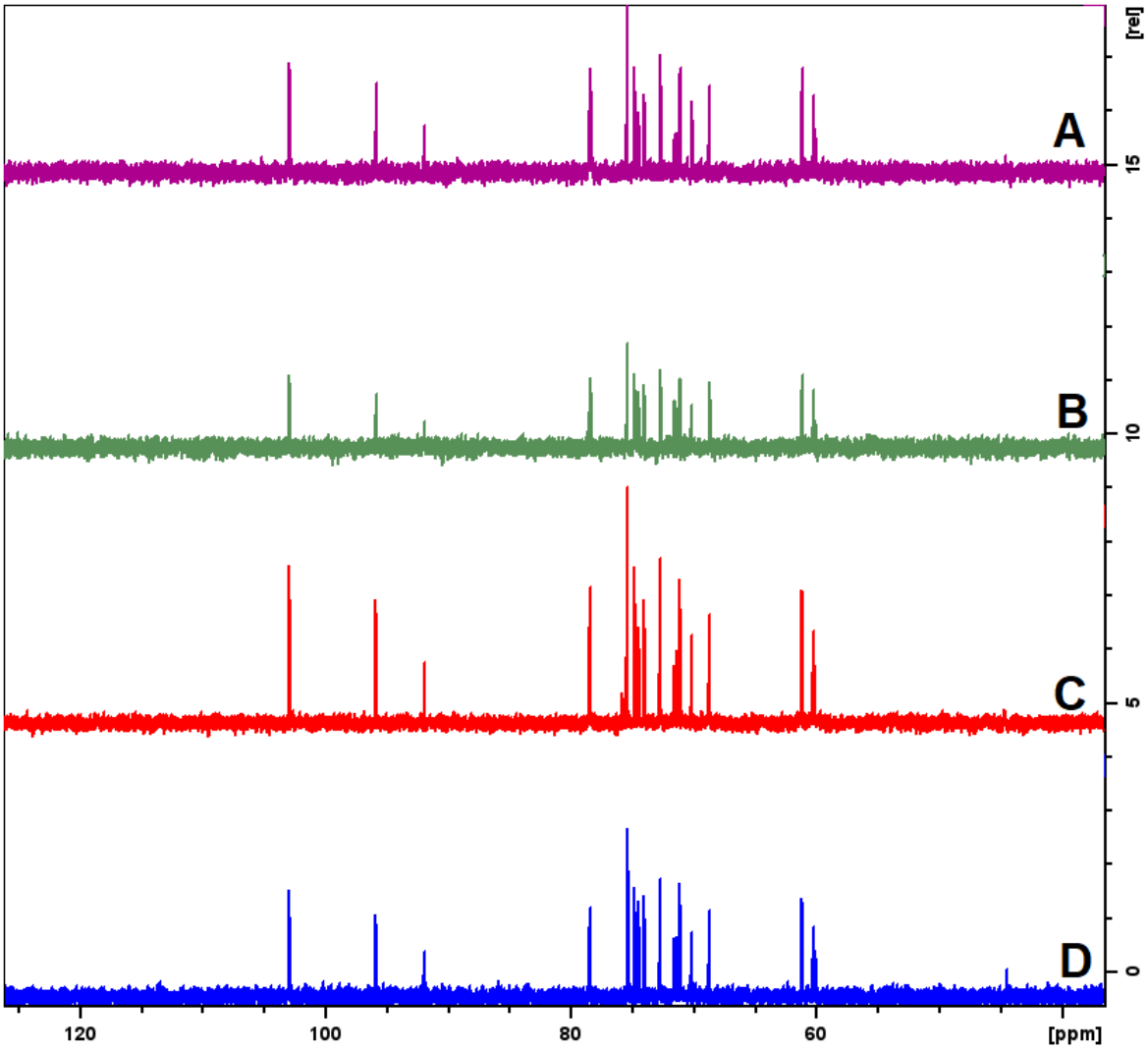


Figure 3. ^{13}C NMR of β-CN-A2 at pH 6.7 (A), β-CN-A2 at pH 4.6 (B), β-CN-A1 at pH 6.7 (C) and β-CN-A1 at pH 4.6 (D).

3.3 Fourier Transform Infrared spectroscopy (FTIR)

The FTIR spectroscopy was used for to observe the presence of variations in the secondary structure of β-CN based on the pH and genetic variants. The C=O stretching vibrations for Amide I (1600-1700 cm^{-1}) and C-N stretching and NH bending for Amide II region (1500-1600 cm^{-1}) were selected for analysis of the structural components in β-CN (Carbonaro & Nucara, 2010; Curley *et al*., 1998). For better peak identification the spectra were curve fitted and shown in Figure 4.

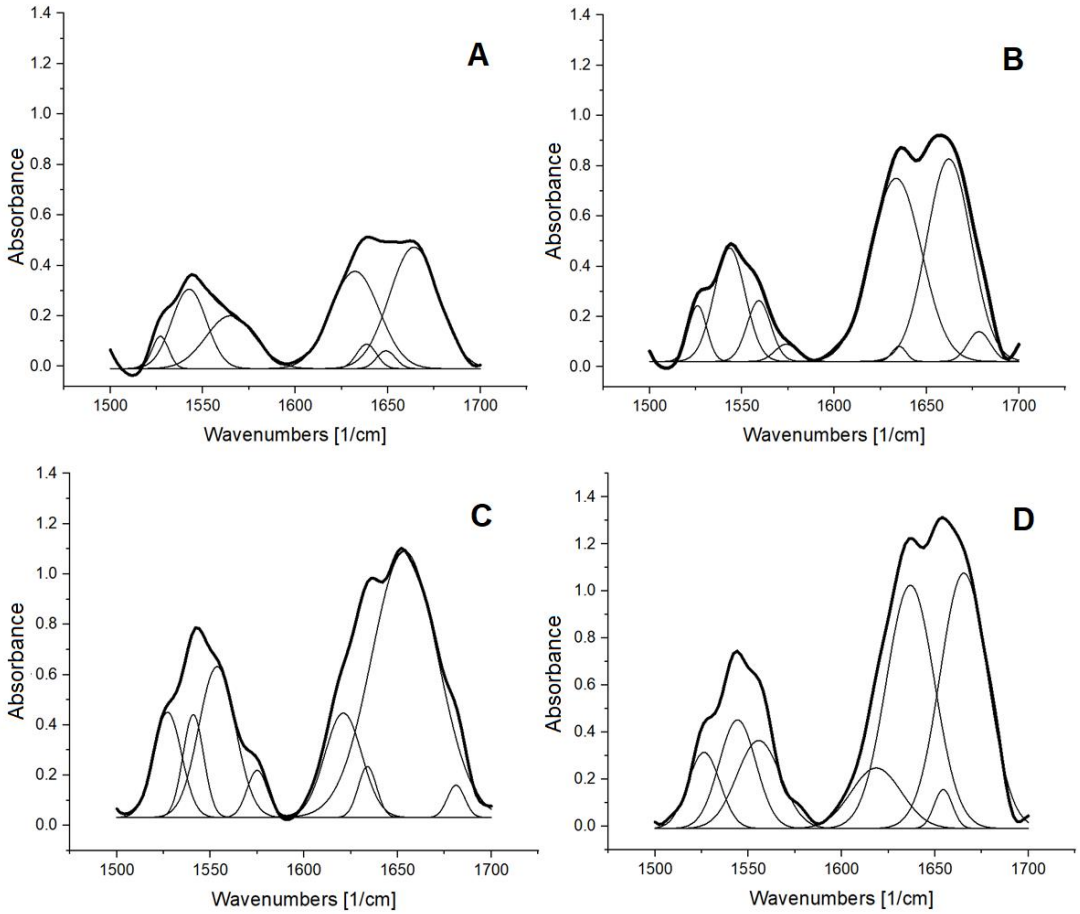


Figure 4. Deconvoluted FTIR spectra in the Amide I and II region of β-CN-A1 at pH 6.7 (A), β-CN-A1 at pH 4.6 (B), β-CN-A2 at pH 6.7 (C) and β-CN-A2 at pH 4.6 (D).

In Amide I region the conformational changes were observed in the variations of peak intensities for the aggregated β-sheets (1700-1680 cm^{-1}), β-turns (1660-1670 cm^{-1}), α-helix (1650-1655 cm^{-1}), random structures (1645 cm^{-1}) and β-sheets (1620-1635 cm^{-1}) (Kher *et al*., 2007). The peak assignment in Amide II region was for turns (1578-1555 cm^{-1}), α-helix (1555-1543 cm^{-1}) and β-sheets (1542-1525 cm^{-1}). The purpose of selection of Amide II region was to provide better fit of Amide I. The spectra absorbance for both β-CN-A1 and β-CN-A2 increased when pH was reduced from 6.7 to 4.6, which was more intensified for β-CN-A1.

In Amide I region of β-CN-A1 at pH 6.7 major peaks were observed at 1635 cm^{-1} and 1670 cm^{-1} corresponding to β-sheets and β-turns, respectively (Figure 4A). Minor peaks were present at 1640 cm^{-1} and 1650 cm^{-1} showing presence of random coil and α-helix, respectively. When pH was changed to 4.6 the α-helix at 1650 cm^{-1} was lost and agregated

β-sheets at 1980 cm^{-1} appeared (Figure 4B). Hence, at pH 4.6, β-CN-A1 molecule promoted aggregation which is expected owing to be the isoelectric point of the casein (Huppertz, 2012). In regard to the Amide II region of the β-CN-A1 at pH 6.7 the major peaks were observed at 1525 cm^{-1}, 1545 cm^{-1}, and 1575 cm^{-1}, corresponding to β-sheet, α-helix and turns, respectively. When pH was altered to 4.6 the peak at 1575 cm^{-1} (turns) was split into two peaks showing presence of new turn structures that can be part of the protein aggregation as observed in Amide I region.

The β-CN-A2 shown completely different pattern of the structural components in the Amide I region (Figure 4C, D). In the Amide region the β-CN-A2 at pH 6.7 had one massive peak at 1670 cm^{-1} (β-turns) and smaller peaks at 1630 cm^{-1} (β-sheets), 1645 cm^{-1} (random coils) and 1685 cm^{-1}(aggregated β-sheets). Thus, no α-helix was observed, however natural aggregated β-sheet structures were present in the molecule. The change of pH to 4.6 (Figure 4D) resulted in high amount of random structures (1645 cm^{-1}) initiated by high volume of aggregated β-sheets. In addition, small peaks were observed for β-sheets (1630 cm^{-1}) and turns (1660 cm^{-1}). The structure of β-CN-A2 was drastically transformed with pH lowered close to its isoelectric point confirming high amount of irregular/random aggregations. The peak at 1645 cm^{-1} or the random coil as observed in this work was previously assigned for polyproline II structures (PPII) (Farrell Jr *et al*., 2001). Moreover, it was previously predicted that β-CN-A2 variant promotes formation of PPII structures that is in agreement with our study (Raynes *et al* 2015). However, in this work it can be observed that the presence of PPII structures is greatly dependent on pH of the solution, which is crucial part of the milk processing and thus important in understanding coagulation differences between A1 and A2 milk. The Amide II region is in agreement with the Amide I spectra for the β-CN-A2 at pH 6.7 (Figure 4C). Thus, the two peaks at 1578 cm^{-1} and 1560 cm^{-1} confirm high amount of turns present in the molecule. The peaks at 1525 cm^{-1} and 1540 cm^{-1} confirms presence of β-sheets, and these two peaks can be related to presence of two forms of β-sheets. At pH 4.6, the β-CN-A2 molecule in the Amide II region has shown presence of three peaks at 1525 cm^{-1}, 1548 cm^{-1} and 1560 cm^{-1}. Moreover, the spectral restructuring has shown that major components affected by the pH lowering to 4.6 for β-CN-A2 were β-sheets and β-turns, which were mainly involved in aggregation as observed in Amide I region.

Principal components analysis was used to identify how distinctive these samples are in regards to the secondary structural components in Amide I and II regions. In Supplementary Figure 3A, the PC loading scores are presented showing 96.9 % variance for PC1 and 2.2% variance for PC2. The separation of the samples was based on the pH. Hence, PC1 discriminated samples at pH 6.7 from the samples at pH 4.6 regardless of the genotype. The pH adjustment had major impact on the structural orientation of the molecules. In the spectral loading plot (Supplementary Figure 3B), the PC1 loading spectra have shown only a presence of loadings at approximately 1685-1700 cm^{-1} for β-CN (A1 and A2) at pH 4.6. Moreover, the presence of aggregaton for samples at pH 4.6 was clearly identifyed by the PC1 loading plot. The PC2 loading plot did not show a significant variance to discriminate samples based on the observed variabilities. The distribution of the loadings used for the PCA analysis is shown in Supplementary Figure 3 C. The loadings confirm presence of defined structural components in the positive end of the PC1 loading score which is for the samples at pH 6.7.

3.4 Raman spectroscopy

The deconvoluted Raman spectrum fitted with Gaussian components in Amide I region (1600-1700 cm^{-1}) of β-CN samples is shown in Figure 5. The band assignments were performed accordingly to the pervious literature (Byler *et al*., 1988: Kuhar *et al*., 2021). Thus, the observed structural features were aromatic side chains (1600-1620 cm^{-1}), β-sheet (1630-1633 cm^{-1}), random coil (~1640 cm^{-1}), α-helix (1646-1655 cm^{-1}), unspecified or turns structures (1661 cm^{-1}), β-sheet (1668-1675 cm^{-1}) and turns (1980-1700 cm^{-1}). In Figure 5A, the bands of the Raman spectrum are shown for β-CN-A1 at pH 6.7. The observed peaks were at 1615 cm^{-1}, 1630 cm^{-1}, 1650 cm^{-1}, 1660 cm^{-1},1680 and 1690 cm^{-1} corresponding to side chains, β-sheets, α-helix, turns, β-sheets and second set of turns, respectively.

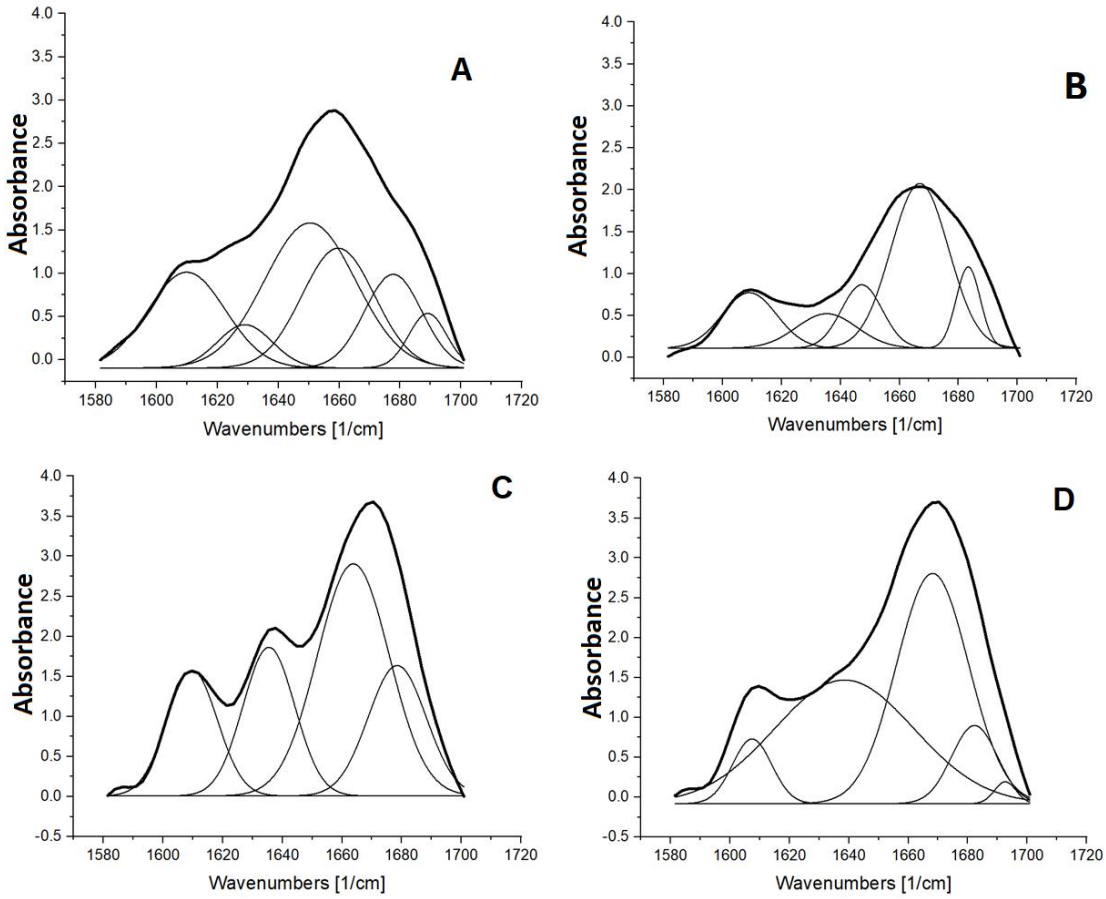


Figure 5. Deconvoluted Raman spectra in the Amide I and II region of β-CN-A1 at pH 6.7 (A), β-CN-A1 at pH 4.6 (B), β-CN-A2 at pH 6.7 (C) and β-CN-A2 at pH 4.6 (D).

Moreover, the β-CN-A1 at pH 6.7 had clearly defined structural features. When pH was reduced to 4.6 the secondary structure rearranged as shown in Figure 5B. The present bands were at 1610 cm^{-1}, 1635 cm^{-1}, 1650 cm^{-1},1672 cm^{-1}, and 1689 cm^{-1} related to aromatic side chains, β-sheets, α-helix, β-sheets and turns, respectively. Hence, pH change from 6.7 to 4.6 affected the α-helix by decrease of peak absorbance, in the same time the β-sheets located at 1678 cm^{-1} promoted band increase. The pH change resulted in the formation of new β-sheet components which are probably result of aggregation behavior of β-CN-A1 at this pH level.

The Raman spectrum of β-CN-A2 at pH 6.7 is shown in Figure 5C. In the deconvoluted spectrum, four major peaks were observed including 1610 cm^{-1}, 1638 cm^{-1}, 1670 cm^{-1}, and 1685 cm^{-1} assigned to side chains, random coils, β-sheets and turns, respectively. In Raman spectra and as previously observed in FTIR spectra results for β-CN-A2 at pH

6.7, the α-helix was not present in the molecule, however, high amount of β-sheets and random coil were dominating.

When pH was altered to 4.6 in β-CN-A2, the observed bands were at 1610 cm^{-1}, 1640 cm^{-1}, 1670 cm^{-1}, 1682 cm^{-1}, and 1695 cm^{-1} corresponding to aromatic side chains, random coils, β-sheets and two sets of turns, respectively. The broad peak at 1640 cm^{-1} was around half of the sum of all peaks in the Amide I region. Thus, the random structures were dominant in β-CN-A2 at pH 6.7 coupled with two sets of turns. The high present of random coils for β-CN-A2 at pH 4.6 was also observed by FTIR spectra. The observed random structures in this part of the spectra (1638-1645 cm^{-1}) can also relate to presence of polyproline II structures (PPII) (Lazarev *et al*., 1985). Both caseins formed different types of aggregates at pH 4.6, which for βCN-A1 was observed to be aggregated β-sheets and α-helix, however for β-CN-A2 the aggregated features had a random coil (PPII helix) distribution. The great amount of PPII in β-CN-A2 was previously confirmed in other studies (Raynes *et al*., 2015).

The PCA analysis separated the samples based on the observed variants (Supplementary Figure 4A and B). In the loading plot, the PC1 shows variance of 61.5 % and sort out the samples along the PC1 axis based on the pH value. Moreover, β-CN (A1 and A2) at pH 6.7 formed individual groups on the positive side of the PC1 axis and β-CN (A1 and A2) with pH 4.5 shown grouping on the negative side of the PC1 axis. The PC2 axis showed variance of 19.9 % and separated the samples based on pH. The PC1 loading plot (Supplementary Figure 4B) did not show significant peak distribution related to the individual groups of the PCA score. Moreover, the low variance for both PC1 and PC2 did not confirm significant differences between the groups along the PC1 and PC2 axis and thus these are not discussed.

3.5 Particle size and microstructure of the particles

The size of the β-CNs particles was observed to confirm how the size of the molecules was affected in regards to the pH effect and genetic variant (Figure 6). Under the native conditions as found in the milk (pH is 6.7), the free β-CNs (not in a micellar state) have different particle size for β-CN-A1 and β-CN-A2. The β-CN in milk have tendency to self-associate (Huppertz, 2013) and form random particles of different sizes. For β-CN-

A2 these particles are bigger than those consisting of β-CN-A1. The pH of 4.6 is the isoelectric point of β-CN and under these conditions, the β-CN is expected to promote aggregation, which would affect the primary structure of the molecule and lead to reformation of the particles (Anema *et al*., 2004). Hence, at pH 4.6 it was observed that the size of the β-CN for both genetic variants (A1 and A2) was reduced more than three times than that under the native conditions (pH 6.7). The trend was the same for both genetic variants.

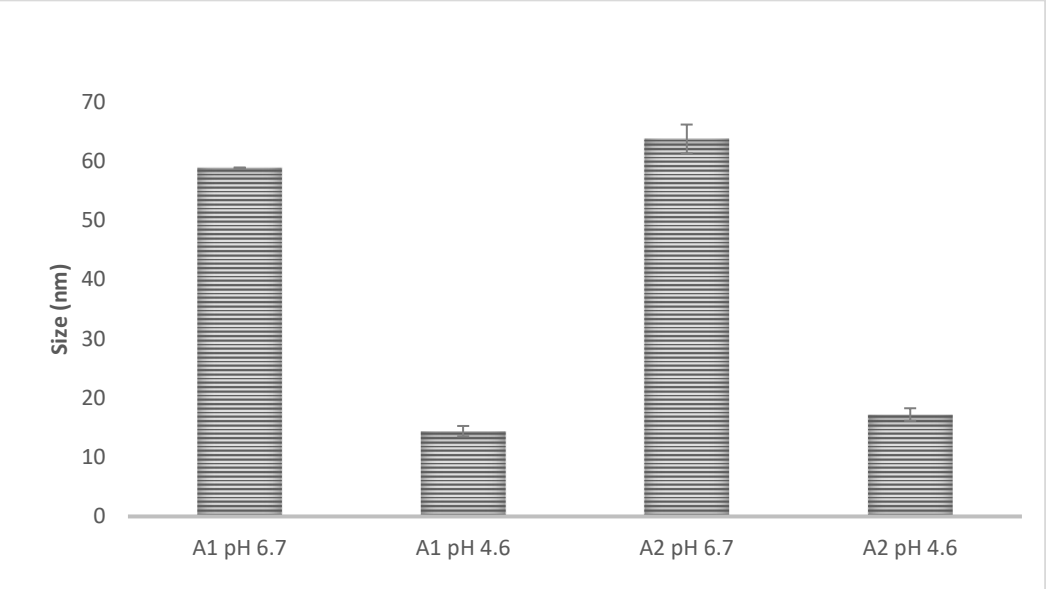


Figure 6. Particle size (nm) of β-CN molecule from A1 and A2 genetic variant and pH of 4.6 and 6.7.

The results from scanning electron microscopy (SEM) have shown differences between the morphology of β-CN from variant A1 and A2 and when the pH was reduced from 6.7 to 4.6. At pH 6.7 the microstructure of β-CN-A1 shows a compact protein network surrounded by small pores (Figure 7A). However, the β-CN-A2 at pH 6.7 shown to have more open protein network distribution enclosed with huge pores (Figure 7C). The pH alteration to 4.6 had significant impact on protein network appearance. Hence, β-CN-A1 formed stiff aggregates with small pores and thin strands (Figure 7B). The β-CN-A2 at pH 4.6 formed open network with large empty cavities.

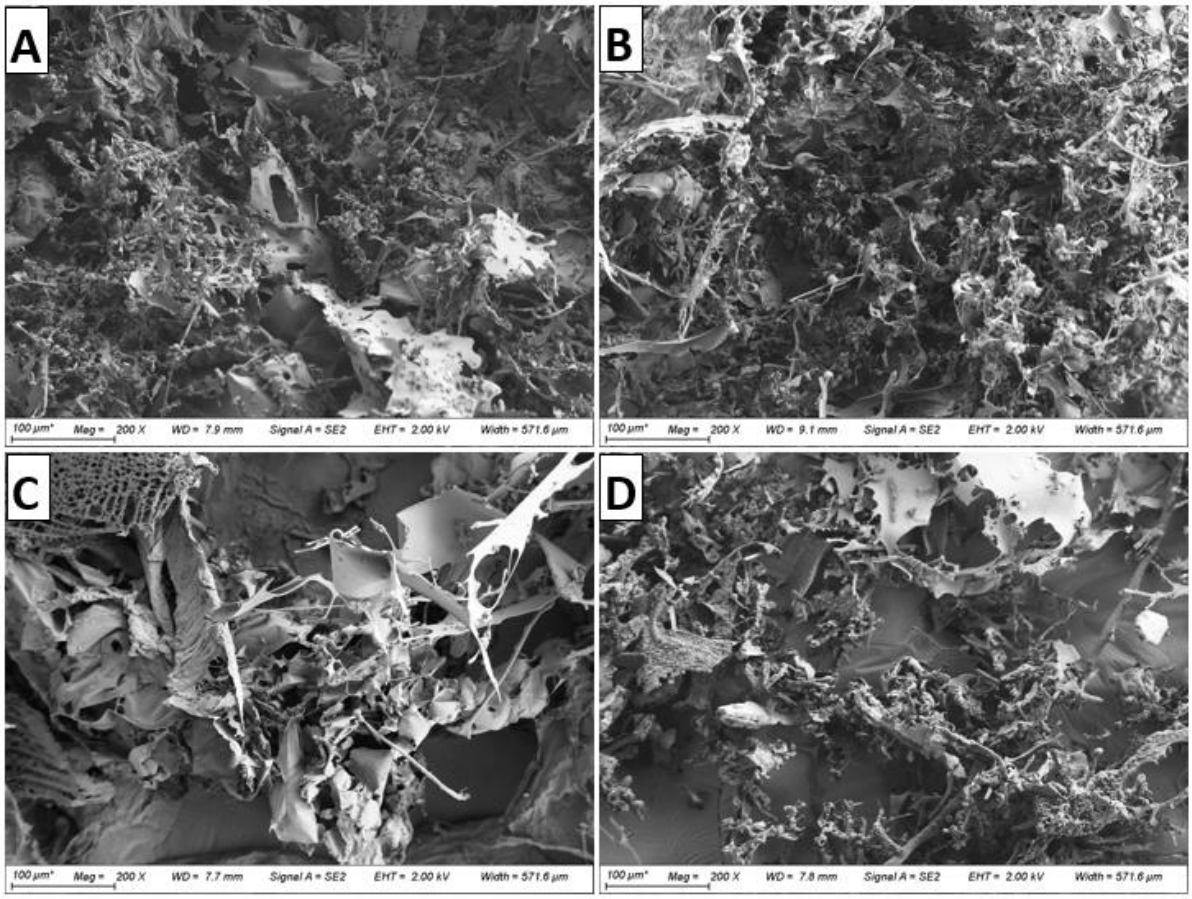


Figure 7. Scanning electron microscopy presenting the microstructure of β-CN-A1 at pH 6.7 (A), β-CN-A1 at pH 4.6 (B), β-CN-A2 at pH 6.7 (C) and β-CN-A2 at pH 4.6. The magnification of images is 200x and scale bar represents 100µm.

The acid-induced aggregates for A1 milk was observed to be dense and more homogeneous in comparison to A2 milk, which exhibit coarse aggregates with large water holding cavities (Daniloski *et al*., 2022). Similarly, in yoghurt produced from A2 milk the existence of highly porous gels with thinner protein strands was observed resulting in lower gel strength in comparison to A1 milk gels (Nguyen *et al*., 2018). Our work is in agreement with the expected coagulation differences existing between A1 and A2 genotype. However, β-CN-A2 at pH 6.7 (which is native pH of milk) was found to exist in large aggregated structures and more open network compared to β-CN-A1. Moreover, before even the aggregation happens the β-CNs from A1 and A2 genotype exist in different structural state. Therefore, we can confirm that the acid induced coagulation differences that exist between A1 and A2 milk are primarily guided by the structural differences in native β-CN molecule.

The observed structural differences and coagulation properties of both genetic variants can affect several characteristics of the final products including texture, melting, elasticity, stretching and many more. The observed aggregates in the β-casein before aggregation affects the formation of the desired gels for productions of milk products as cheese and yogurt. Moreover, not only the quality of the product is affected by the differences in the formed gels between the both variants but also variations in yield and shelf life would be expected.

4. Conclusion

In this work, it was observed how the β-CN-A1 and β-CN-A2 differ in regards to the structural characteristics at pH of 6.7 and 4.6. Moreover, the selected pH is crucial for the native conditions in which the caseins are found in milk (pH 6.7) and conditions that are used for milk coagulation while producing fermented milk products (pH 4.6). The combination of the methods used in this study successfully discriminated the structural differences between β-CN-A1 and β-CN-A2 samples. FTIR and Raman spectroscopy presented existence of different structural components for both genotypes regardless the pH. The microstructure and the size of the particles determined the protein network distribution and its size. The open and porous network in native and aggregated β-CN-A2 was due to presence of irregular structures and high presence of PPII. Moreover, the aggregation because of pH was more intensified for β-CN-A1 as observed in ^{13}C NMR spectra and resulted with aggregated β-sheets that build up dense protein network. The findings in this study assist in understanding the impact of β-CN for both genotypes (A1 and A2) on the coagulation properties of A1 and A2 milk. Moreover, this study for the first time reveals the structural differences of both β-CN-A1 and β-CN-A2 in conditions in which these caseins are found in milk and while coagulated for yoghurt or cheese production.

References

Anema, S.G., Lowe, E.K. and Lee, S.K., 2004. Effect of pH at heating on the acid-induced aggregation of casein micelles in reconstituted skim milk. *LWT-Food Science and Technology*, *37*(7), pp.779-787.

Byler, D.M., Farrell Jr, H.M. and Susi, H., 1988. Raman spectroscopic study of casein structure. *Journal of Dairy Science*, *71*(10), pp.2622-2629.

Carbonaro, M. and Nucara, A., 2010. Secondary structure of food proteins by Fourier Transform spectroscopy in the mid-infrared region. *Amino acids*, *38*(3), pp.679-690.

Curley, D. M., Kumosinski, T. F., Unruh, J. J., & Farrell, H. M. (1998). Changes in the secondary structure of bovine casein by Fourier Transform Infrared Spectroscopy: Effects of calcium and temperature. *Journal of Dairy Science, 81*(12), 3154-3162.

Daniloski, D., McCarthy, N.A., Gazi, I. and Vasiljevic, T., 2022. Rheological and structural properties of acid-induced milk gels as a function of β-casein phenotype. *Food Hydrocolloids*, *131*, p.107846.

Daniloski, D., McCarthy, N.A., O'Callaghan, T.F. and Vasiljevic, T., 2022. Authentication of β-casein milk phenotypes using FTIR spectroscopy. *International Dairy Journal*, *129*, p.105350.

Daniloski, D., Cunha, N.M., McCarthy, N.A., O'Callaghan, T.F., McParland, S. and Vasiljevic, T., 2021. Health-related outcomes of genetic polymorphism of bovine β-casein variants: A systematic review of randomised controlled trials. *Trends in Food Science & Technology*, *111*, pp.233-248.

Dong, X., Gong, Z., Lu, Y.B., Liu, K., Qin, L.Y., Ran, M.L., Zhang, C.L., Liu, Z., Zhang, W.P. and Tang, C., 2017. Ubiquitin S65 phosphorylation engenders a pH-sensitive conformational switch. *Proceedings of the National Academy of Sciences*, *114*(26), pp.6770-6775.

Farrell Jr, H.M., Wickham, E.D., Unruh, J.J., Qi, P.X. and Hoagland, P.D., 2001. Secondary structural studies of bovine caseins: temperature dependence of β-casein

structure as analyzed by circular dichroism and FTIR spectroscopy and correlation with micellization. *Food Hydrocolloids*, *15*(4-6), pp.341-354.

Huppertz, T., 2012. Chemistry of the caseins. In *Advanced Dairy Chemistry: Volume 1A: Proteins: Basic Aspects, 4th Edition* (pp. 135-160). Boston, MA: Springer US.

Huppertz, T., Hennebel, J.B., Considine, T., Kelly, A.L. and Fox, P.F., 2006. A method for the large-scale isolation of β-casein. *Food chemistry*, *99*(1), pp.45-50.

Illa, O., Olivares, J.A., Nolis, P. and Ortuno, R.M., 2017. The relevance of the relative configuration in the folding of hybrid peptides containing β-cyclobutane amino acids and γ-amino-L-proline residues. *Tetrahedron*, *73*(44), pp.6286-6295.

Ivanova, G., Yakimova, B., Angelova, S., Stoineva, I. and Enchev, V., 2010. Influence of pH on the *cis–trans* isomerization of Valine-Proline dipeptide: An integrated NMR and theoretical investigation. *Journal of Molecular Structure*, *975*(1-3), pp.330-334.

Kher, A., Udabage, P., McKinnon, I., McNaughton, D. and Augustin, M.A., 2007. FTIR investigation of spray-dried milk protein concentrate powders. *Vibrational Spectroscopy*, *44*(2), pp.375-381.

Kleckner, I.R. and Foster, M.P., 2011. An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1814*(8), pp.942-968

Kuhar, N., Sil, S. and Umapathy, S., 2021. Potential of Raman spectroscopic techniques to study proteins. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *258*, p.119712.

Lazarev, Y.A., Grishkovsky, B.A. and Khromova, T.B., 1985. Amide I band of IR spectrum and structure of collagen and related polypeptides. *Biopolymers: Original Research on Biomolecules*, *24*(8), pp.1449-1478.

Lucey, J.A., 2002. Formation and physical properties of milk protein gels. *Journal of dairy science*, *85*(2), pp.281-294.

Markoska, T., Daniloski, D., Vasiljevic, T., Orbell, J., and Huppertz, T., 2022. Temperature- and pH-induced structural changes of β-casomorphin 11-A1 and -A2 studied by Nuclear Magnetic Resonance, Fourier-Transform infrared spectroscopy, chemometrics and molecular modelling. *Current research in Food Science- under review*.

Markoska, T., Daniloski, D., Vasiljevic, T. and Huppertz, T., 2021. Structural changes of β-casein induced by temperature and pH analysed by Nuclear Magnetic Resonance, Fourier-Transform infrared spectroscopy, and chemometrics. *Molecules*, *26*(24), p.7650.

Mediwaththe, A., Chandrapala, J. and Vasiljevic, T., 2018. Shear-induced behaviour of native milk proteins heated at temperatures above 80° C. *International dairy journal*, *77*, pp.29-37.

Nguyen, H.T., Schwendel, H., Harland, D. and Day, L., 2018. Differences in the yoghurt gel microstructure and physicochemical properties of bovine milk containing A1A1 and A2A2 β-casein phenotypes. *Food Research International*, *112*, pp.217-224.

Poulsen, N.A., Bertelsen, H.P., Jensen, H.B., Gustavsson, F., Glantz, M., Månsson, H.L., Andrén, A., Paulsson, M., Bendixen, C., Buitenhuis, A.J. and Larsen, L.B., 2013. The occurrence of noncoagulating milk and the association of bovine milk coagulation properties with genetic variants of the caseins in 3 Scandinavian dairy breeds. *Journal of Dairy Science*, *96*(8), pp.4830-4842.

Raynes, J.K., Day, L., Augustin, M.A. and Carver, J.A., 2015. Structural differences between bovine A1 and A2 β-casein alter micelle self-assembly and influence molecular chaperone activity. *Journal of dairy science*, *98*(4), pp.2172-2182.

Sah, B.N.P., Vasiljevic, T., McKechnie, S. and Donkor, O.N., 2016. Physicochemical, textural and rheological properties of probiotic yogurt fortified with fibre-rich pineapple peel powder during refrigerated storage. *LWT-Food Science and Technology*, *65*, pp.978-986.

Seavey, B.R., Farr, E.A., Westler, W.M. and Markley, J.L., 1991. A relational database for sequence-specific protein NMR data. *Journal of Biomolecular NMR*, *1*, pp.217-236.

Silverstein, R.M. and Bassler, G.C., 1962. Spectrometric identification of organic compounds. *Journal of Chemical Education*, *39*(11), p.546.

Zhang, Y., Liu, D., Liu, X., Hang, F., Zhou, P., Zhao, J., Zhang, H. and Chen, W., 2018. Effect of temperature on casein micelle composition and gelation of bovine milk. *International Dairy Journal*, *78*, pp.20-27.

Supplementary material for: “Diversities between β-CN from A1 and A2 milk in regards to the pH and genetic alternative“

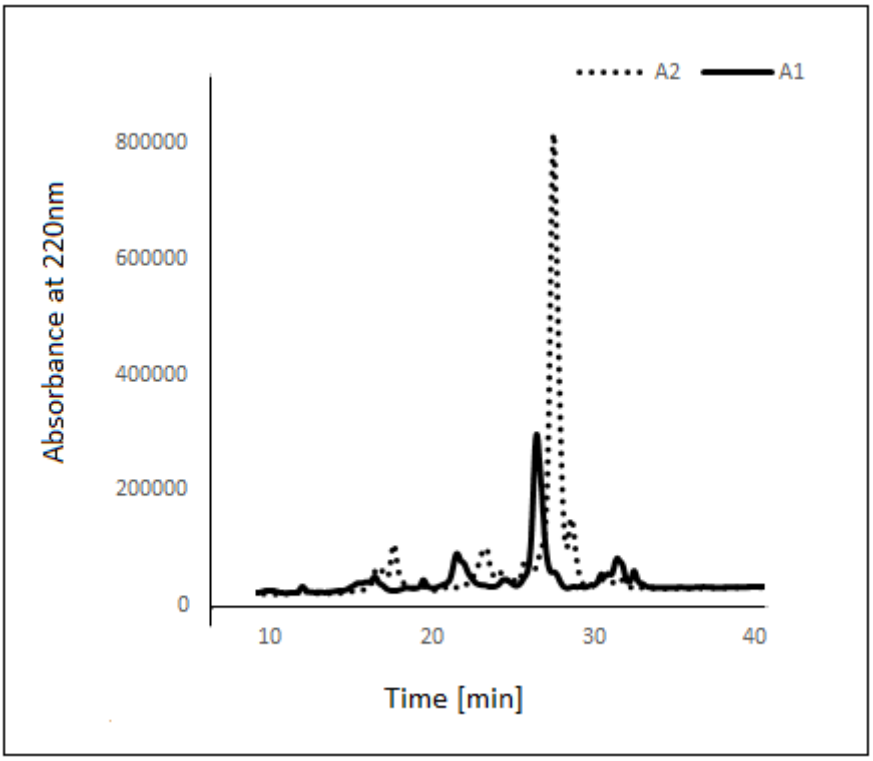


Figure S1. RP-HPLC chromatographic profile for detection of β-CN variant. The dotted peak at 26.84 min is assigned to A2 β-CN and the full line peak at 25.64 min is assigned to A1 β-CN.

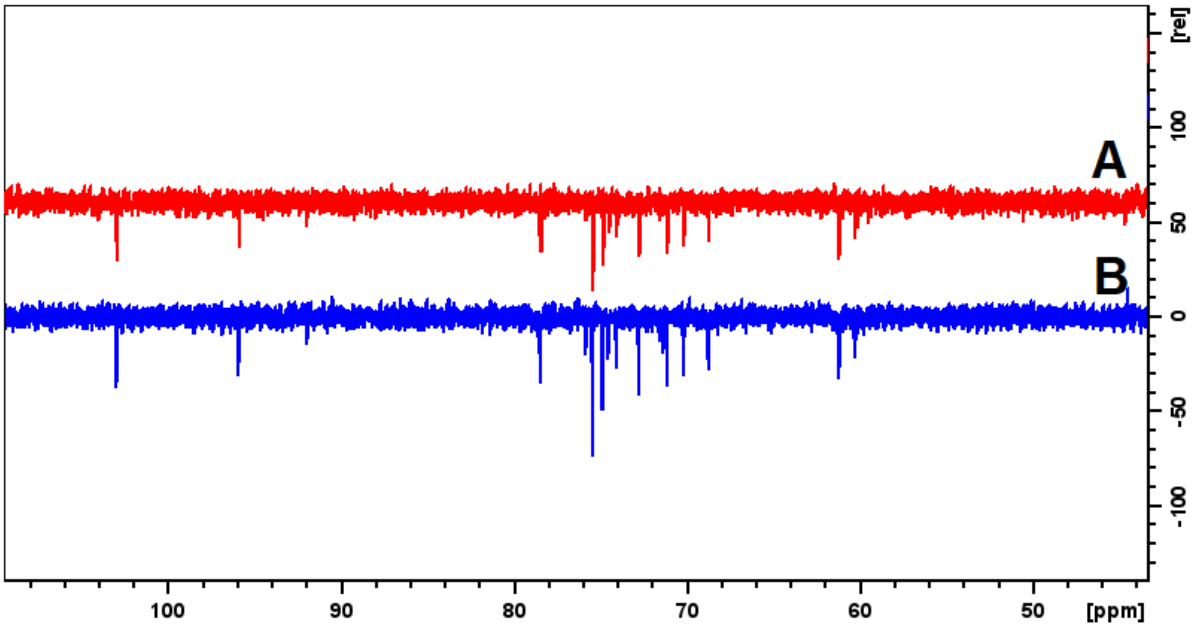


Figure S2. ^{13}C NMR spectra calculated differences between β-CN-A2 at pH 6.7 and 4.6 (A) and spectra difference of β-CN-A1 at pH 6.7 and 4.6 (B)

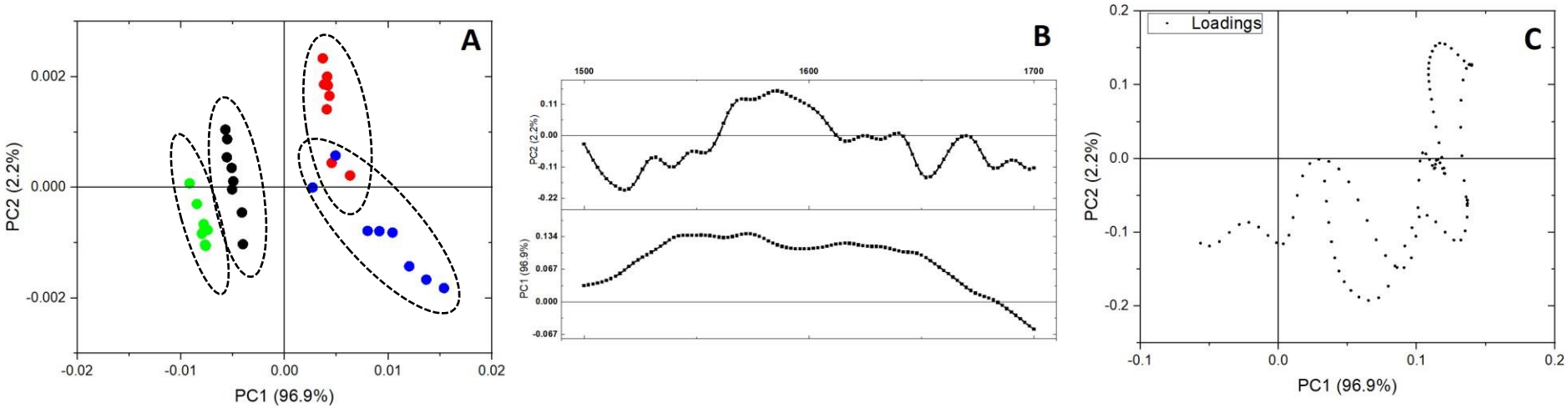


Figure S3. Principal component analysis of Amide I and II for FTIR spectra. In the score plot (A) the scores in black are for β-CN-A1 at pH 4.6, in red β-CN-A1 at pH 6.7, in blue β-CN-A2 at pH 6.7 and in green β-CN-A2 at pH 4.6; The loading plot spectra for PC1 and PC2 is shown in B: The distribution of the loadings is presented in C.

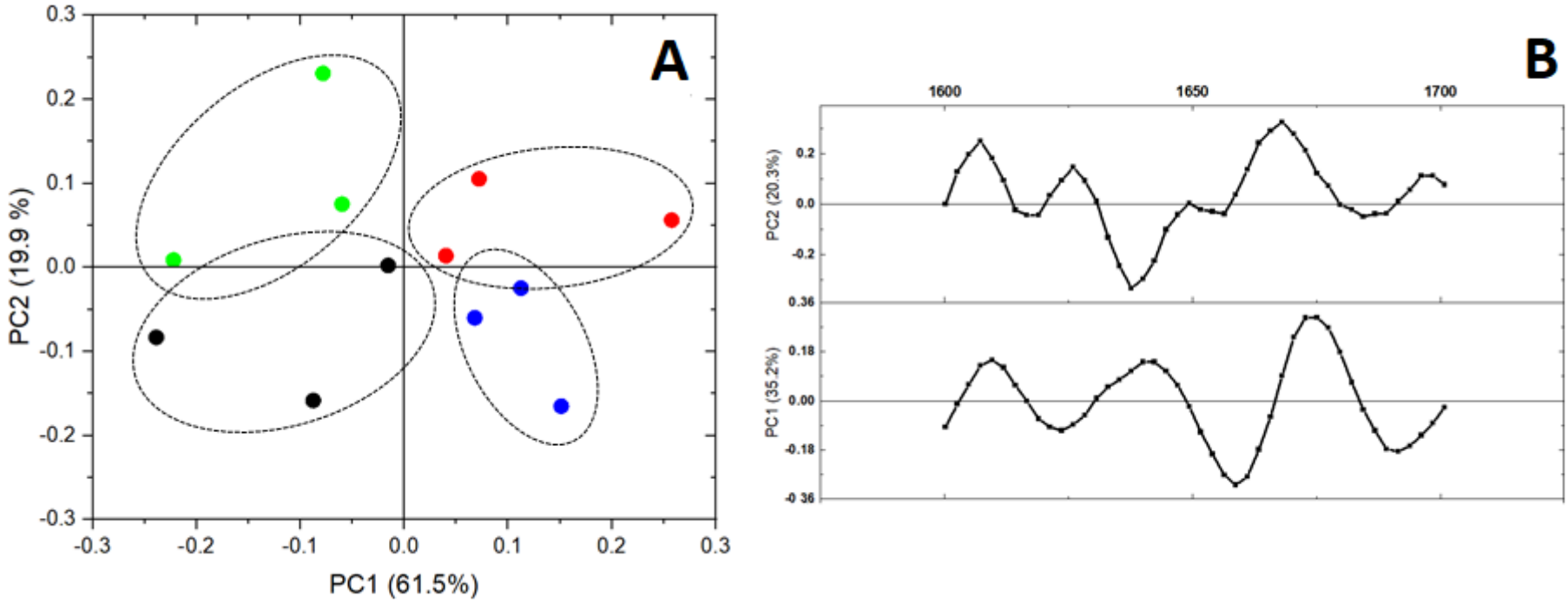


Figure S4. Principal component analysis of the second derivative of Amide I for Raman spectra. In the score plot (A) the scores in black are for β-CN-A1 at pH 4.6, in red β-CN-A1 at pH 6.7, in blue β-CN-A2 at pH 6.7 and in green β-CN-A2 at pH 4.6; The loading plot spectra for PC1 and PC2 is shown in B.

CHAPTER 7: Conclusions and future directions

7.1 Overall conclusions

Understanding the structural changes of milk proteins during processing is essential for securing high quality control during milk manufacturing. Moreover, the structure of milk proteins can also affect the digestibility and hence liberalization of opioid peptides in the human gut. Nuclear Magnetic Resonance (NMR) and Fourier Transform Infrared (FTIR) spectroscopy as advanced analytical tools have contributed significantly in bringing a new knowledge to the dairy science. Moreover, NMR can provide greater understanding at a molecular level by identifying the position of individual protons. This observation can assist in understanding the possible isomerisms, binding affinities and conformational orientations at a molecular level. On the other hand, FTIR offers insights into secondary structure orientation by identifying presence of defined structures including α-helix, β-sheets or turns. The combination of these two analytical tools can successfully provide deep knowledge of structural characteristics of milk components under native and modified conditions.

The BCM7 as most present opioid peptide liberated from A1 and/or A2 β-casein is characterised with different *cis-trans* isomerism as a result of different pH environments presenting predominant *trans* conformers and increase of *cis* isomerism at pH 6.7. The structural changes were observed to be dominant and distinguishing at Gly^{5}-Pro^{6} bond. The enzymes that cleave large molecules into smaller fragments can have different activity depending on the pH environment and availability of the free peptide bonds. The activity and the occurrence of free BCMs in the human body can be strongly dependent on the *cis-trans* isomerism of X-Pro. The isomerism of the X-Pro bond can have great effect on peptide folding and conformations and thus can affect protein function and biological pathway. BCM7 peptide is predominately found released during digestion of A1 β-casein variant. The preferred enzyme hydrolysis of β-casein A1 variant can be due to different isomeric form of BCM7 when found in the β-casein molecule compare to β-casein from A2 variant. Hence, analysing larger peptides containing the distinguishing pattern of A1 and A2 β-casein (BCM11) can better individualise the main structural differences of the peptides and their effect on milk processing.

The BCM11 from A1 β-casein have more flexible and loose conformation. This flexibility appears to be due tautomeric behaviour of His^8 (imidazole ring) governed by pH and temperature variations. The Ile^7-His^8 bond had high exposure to the solvent. However, the BCM11 from A2 β-casein have more packed structure whereas the Ile^7-Pro^8 is closed in the internal structure of the molecule. The imidazole from His^8 at high pH and temperature variations had rapid proton exchangeability with the water or was involved in hydrogen bonding. Thus, the imidazole ring of His from BCM11 from A1 β-casein has unstable conformation in regards to pH and temperature changes compared to BCM11-A2 peptide. The *cis-trans* isomerism has shown rapid exchangeability between *cis* and *trans* isomers due to high flexibility. Moreover, the *trans* isomers were dominant for BCM11 from A2 β-casein and *cis* for BCM11 from A1 β-casein. The conformational properties of BCM11 are important for the enzyme affinity and accessibility for peptide bond cleavage that is found to be the major factor in the discussions for the health benefits of A1/A2 milk.

Both pH and temperature variations are crucial factors in milk processing and human digestion. For β-casein, temperature of 4 °C promotes lowering of hydrophobicity and molecular unfolding to the native conformers and thus dissociation of the monomers from the casein micelle. The FTIR spectroscopy depicted the restructuring of the β-casein by confirming great presence of random coils at 4 °C at the expense of the α-helical structure. The increase of pH to neutral had large effect on the monomer transition. Both NMR and FTIR spectroscopy detected structural transition of β-casein when pH and temperature were modified confirming high structural flexibility of the molecule. The structural differences of A1 and A2 β-casein leads to formation of strong and compact gels for A1 β-casein and gels with open and porous network for A2 β-casein. The variations in the gel formation was also shown by larger particles observed in A2 β-casein compared to A1 β-casein. The structural distribution of A1 and A2 β-casein variants can be a crucial factor in detecting the processing differences occurring when dealing with A1 and A2 milk. This includes lower aggregation properties of A2 milk compared to A1 milk. The structural differences of A1 and A2 β-casein are the major factor that contribute to lower favourability of A2 milk in production of cheese and yoghurt for industries.

NMR, FTIR and Raman spectroscopy revealed existence of different structural components for both genotypes regardless of the pH. Thus, these genotypes adapt a

secondary structure that further is affected by the processing conditions in different manners and thus resulting in products with distinctive properties. The protein microstructure of β-casein from A1 milk adapts compactly build protein network, which when coagulated reforms dense gels. On the other hand, the β-casein from A2 milk have native aggregates with open structure and large cavities. When coagulated the protein network is composed of irregular gels with massive pores. The findings are crucial for understanding the coagulation properties of A1 and A2 milk.

The combination of spectroscopic tools and chemometrics provided valuable information for identifying the major structural variations between A1 and A2 β-casein and its fragments. The combination of the studies in this project led to significant outcome by detecting the structural variations and molecular transition of β-casein from A1 and A2 milk or its fragments confirming that the genotype has major impact on the structure and thus coagulation properties of these molecules. Moreover, one amino acid (Pro or His) was shown to have substantial effect on different coagulation properties of A1 and A2 milk.

7.2 Future perspective

The A1 and A2 milk genotypes are becoming highly important topic in the market and research. Thus, the differences in the technological and health implications of both genotypes leads to different favourability for industries and consumers. The proven technological issues in production of milk products including cheese and yogurt for A2 genotype is highly unfavourable for dairy industry. However, from consumer aspect A2 was observed to be more beneficial for human health and thus more desired milk variant on the market. In order to comply with both requirements (industry and consumers) the alternative solutions can be found by understanding the basic molecular differences that makes these genetic variants so unique and distinctive. Since β-casein is the major factor that makes the A1 and A2 milk to have different technological properties, its structural adaptability should be considered as factor for adaptability of both genetic variants to industry and consumer needs. Thus, by knowing the structure of the β-casein and β-casomorphins can assist in providing possible solutions in reducing the release of opioid

peptides during digestions and improving the coagulation properties of β-casein that is dependant of the genotype. Future studies are needed to investigate in finding alternative approach in formation of more compact gels for A2 milk genotype and in controlling the enzymatic cleavage of Ile-His peptide bond in A1 β-casein.

In the present study, only a few conditions were selected to observe the structural transitioning of peptides and β-casein from A1 and A2 genotype. Hence, the selected conditions were used to imitate the native conditions in which β-casein is found in milk and conditions that are primarily used for milk processing and/or human digestion. Even though the obtained study conceive vital information of the major differences between the A1 and A2 β-casein genotype it will be worthwhile to observe how the β-casein or it smaller fragments (β-casomorphins) will behave in the presence of enzymes. Hence, the observed differences in the bond accessibility and thus possible different level of peptide cleavage can be clearly identified if both β-caseins and their fragments are introduced with enzymatic reaction.

The methodology used in this study successfully identified transitional changes in the structure of β-casein from A1 and A2 milk. However, the methodology used in this work cannot be completely comparable to the complex digestive system in the human body and the processing conditions applied on the milk systems during production of fermented milk products. Moreover, obtaining comprehensive knowledge on β-casein aggregation during *in vivo* digestion and/or milk processing in industry scale would further identify the structural behaviour of this casein from both genotypes (A1 and A2). This will include employment of different digestive enzymes to interact with the protein and induce *in vivo* digestion. The involvement of enzymes will aim to simulate human digestion process. This will include pepsin and pancreatic to observe the hydrolysis and the level of soluble protein fragments or free amino acids. Additionally, involvement of trypsin, chymotrypsin, peptidase and pronase would be valuable to observe the hydrolysis of the proteins and observe each step of the enzymatic reaction similarly as gastro-intestinal digestion.