

1-1-2011

Structure and Vitamin D₃-Binding Affinity of Milk Proteins

Muhammad Ali Naqvi
Ryerson University

Follow this and additional works at: <http://digitalcommons.ryerson.ca/dissertations>

 Part of the [Cellular and Molecular Physiology Commons](#)

Recommended Citation

Naqvi, Muhammad Ali, "Structure and Vitamin D₃-Binding Affinity of Milk Proteins" (2011). *Theses and dissertations*. Paper 1441.

This Thesis is brought to you for free and open access by Digital Commons @ Ryerson. It has been accepted for inclusion in Theses and dissertations by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact bcameron@ryerson.ca.

STRUCTURE AND VITAMIN D₃-BINDING AFFINITY OF MILK PROTEINS

by

Muhammad Ali Naqvi

B.Sc. Chemistry and Biology

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree

Master of Science

in the Program of

Molecular Science

Toronto, Ontario, Canada, 2011

© Muhammad Ali Naqvi, 2011

Author's Declaration Page

I hereby declare that I am the sole author of this thesis or dissertation. I authorize Ryerson University to lend this thesis or dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis or dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Thesis Abstract

STRUCTURE AND VITAMIN D₃-BINDING AFFINITY OF MILK PROTEINS

Muhammad Ali Naqvi

Master of Science, Molecular Science

Department of Chemistry and Biology

Ryerson University

2011

Two projects formed the basis of this thesis related to protein-vitamin D₃ (VD₃) binding mechanisms and efficacy. First, a method to characterize the binding of VD₃ to food-related macromolecules that may be used for enrichment of milk was devised. Results suggested that sodium caseinate and hydroxylpropyl- β -cyclodextrin could effectively bind VD₃ and may be used as carriers. Secondly, molecular dynamics simulations were used to determine the conformation ensemble of the β -casein phosphopeptide (β -CPP). Radius of gyration, H-bonding, Ramachandran plot, and secondary structure were ascertained, and showed good agreement with simulations of other disordered peptides as well as experimental data of β -CPP. Overall, this new binding assay now affords the ability to study interactions between macromolecules and vitamin D₃. As well, by performing simulations of a single casein peptide, important data needed to understand the intramolecular interactions and structure of β -casein (as well as the casein micelle) have been elucidated.

Acknowledgements

This degree has been a life-changing experience owing to my family, friends, students, mentors. These individuals helped shape me a scientist I am now, through their questions, encouragement, sound advice, criticism, sometimes knowingly and sometimes not. I understand that this list may fall short, yet I will attempt to mention those who have left the most memorable impressions in my life over the course of this degree.

First of all I would like to thank my wife Waufa, my greatest supporter, my best friend, and counsellor, whose smile kept me sane and whose cooking kept me healthy. Secondly, I want to extend my gratitude to my parents, Dr. Hassan Nawab and Jahan Ara Naqvi, who have devoted their life towards raising me and also encouraging me to pursue science. My brothers and sisters who are always make me feel special yet are humble about their own range of talents and my nieces and nephew, Eshal, Eiliya, Zainab, Hussain and Hasnain, whose presence is a constant source of joy.

I would also like to thank Professor D  rick Rousseau, to whom I owe many things, amongst those an introduction to Molecular Science. His far-sightedness saw some potential in this humble student, and nurtured it; all the while making sure work did not seem like a chore. I am truly indebted to him for his support over the years I spent under his supervision. Similarly, I would like to thank Dr. Darrick Heyd, whose unmatched command of multidisciplinary wisdom leaves me striving for more knowledge. He has always been able to provide guidance in difficult questions and has asked some tough ones himself to nudge me over towards the answer. Dr. Christopher Evans was also tremendously influential in my academic development. His depth of knowledge in fluorescence spectroscopy eased my research work considerably. I would also like to thank Dr. David Pink who introduced me along with his student Jordan Marsh, to Molecular Dynamics and Monte Carlo simulation. Development of Molecular Dynamics techniques could not have been possible without the help of Dr. R  gis Pom  s and his PhD student Sarah Rauscher. These two individuals catalyzed my thesis by helping me gain understanding about intrinsically disordered proteins and about Molecular Dynamics methodology.

Of course I couldn't have progressed much without people to show me the ropes. I would like to thank everyone in the Rousseau Food Lab whom I have worked with, and learned from. A special thanks to Sofija Katic from Dr. Evan's lab whose attention to detail and severely analytical technique helped guide me straight through my experimentation.

This list is incomplete without the mention of the faculty and staff of Department of Chemistry and Biology who have been absolutely wonderful. To them and those I have failed to mention, I extend my deepest gratitude.

Dedication

I dedicate this work to my grandfather, my first science teacher, may he rest in peace.

Table of Contents

Thesis Abstract	v
Acknowledgements	vi
Dedication	vii
List of Tables	xi
List of Figures.....	xii
 CHAPTER 1: INTRODUCTION TO VITAMIN D₃ & MILK PROTEINS.....	1
1.1 Focus of the thesis.....	1
1.2 History of the thesis	5
1.3 Aim of the thesis	7
 CHAPTER 2: STRUCTURE & FUNCTION OF VITAMIN D₃ AND MILK PROTEINS... 	8
2.1 Vitamin D₃	8
2.1.1 Nutrition.....	10
2.1.2 Barriers to adequate levels of VD ₃	11
2.1.3 Sources.....	11
2.1.4 Fortification of dairy products	12
2.2 Milk proteins	14
2.3 Whey proteins.....	14
2.3.1 Structure of β -lactoglobulin	14
2.3.2 Function of β -lactoglobulin	15
2.3.3 Ligand binding.....	16
2.4 Caseins	16
2.4.1 Structure	16
2.4.2 Function	18
 CHAPTER 3: BINDING OF VITAMIN D₃ TO FOOD-BASED MACROMOLECULES ..	21
3.1 Abstract.....	21
3.2 Introduction	22
3.2.1 Necessity for an alternative assay to fluorescence	22
3.2.2 Selection of test molecules	24
3.3 Materials and methods	26

3.4 Observations	28
3.4.1 Experimental remarks	28
3.4.2 Assay limits.....	30
3.4.3 Assay efficiency	30
3.5 Results and discussion	32
3.5.1 Kinetics	32
3.5.2 Water and phosphate buffer	33
3.5.3 Whey protein isolate (native/denatured)	34
3.5.4 Cyclodextrin	36
3.5.5 Sodium caseinate	36
3.5.6 Polysorbate 80	40
3.6 Conclusions	40
3.7 Acknowledgements	41
 CHAPTER 4: STRUCTURE OF B-CASEIN F(1-25).....	 42
4.1 Abstract.....	42
4.2 Introduction	43
4.3 Materials and methods:	47
4.3.1 Molecular dynamics Set-up.....	47
4.3.2 Quality assurance.....	48
4.3.3 Simulation analysis.....	48
4.4 Results and discussion	49
4.4.1 Ensemble equilibration	49
4.4.2 Simulation performance	50
4.4.3 Radius of gyration	52
4.4.4 Hydrogen Bonding	55
4.4.5 Secondary Structure.....	59
4.4.5.1 Helix:	62
4.4.5.2 β -Structure (Bridges, Ladders, and Sheets)	62
4.4.5.3 Loops:	63
4.4.5.4 Bends:	63
4.4.5.5 Turns:	64
4.4.6 Ramachandran Angles	65
4.4.7 Working Hypothesis for interaction with calcium and phosphate ions	68
4.5 Conclusions	70
4.5.1 Future Directions	70
4.6 Acknowledgements	72

CHAPTER 5: OVERALL REMARKS.....	73
REFERENCES.....	74

List of Tables

Table 2.1 Structure Function correlation of the Casein Micelle	19
Table 4.1: DSSP representation of β -CPP secondary structure.	45
Table 4.2 DSSP characterization of secondary structures, mechanism and molecular structure.....	60

List of Figures

<u>Figure 1.1</u> History of M.Sc. Thesis.....	5
<u>Figure 2.1</u> Chemical Structure of Vitamin D ₂ . Vitamin D ₃ does not have the double bond and methyl group indicated in red.	8
<u>Figure 2.2</u> History of (A) milk and (B) cheese consumption in Canada.	12
<u>Figure 2.3</u> Crystal Structure of β -lactoglobulin adapted from the RSCB database	15
<u>Figure 2.4</u> Primary sequences of caseins reported from N \rightarrow C terminus, with corresponding chain length, Mw, number of phosphate centers and their loci.	17
<u>Figure 2.5</u> Hypothesized model of the casein micelle with calcium phosphate nanoclusters.....	18
<u>Figure 3.1</u> Vitamin D ₃ Solubility Assay with step-wise pictures from P80 samples and 10mg of VD ₃ layer	28
<u>Figure 3.2</u> Standard Curve of VD ₃ in EtOH	30
<u>Figure 3.3</u> Efficiency of Indirect VD ₃ Assay.....	31
<u>Figure 3.4</u> Kinetic solubility of VD ₃ in 10g/L SCN-PBS solution.	32
<u>Figure 3.5</u> VD ₃ -Macromolecule solubility	33
<u>Figure 3.6</u> Particle size of PBS equilibration with VD ₃ . N=4. Measurements of PBS-only yielded a flat line at 0 intensity at all particle sizes.	34
<u>Figure 3.7</u> VD ₃ -Macromolecule Binding at 10g/L MM and 1mg of VD ₃	35
<u>Figure 3.8</u> Sodium Caseinate micelles (10g/L SCN) before and after equilibration with VD ₃ . N=4.	37
<u>Figure 3.9</u> Plot of ζ -potential SCN before and after equilibration with VD. N=4x3.....	39
<u>Figure 3.10</u> P80 micelles (10g/L) before and after equilibration with VD. N=4.....	40
<u>Figure 4.1</u> Radius of gyration for 100 replica series over time.	49
<u>Figure 4.7</u> Theoretical divisions of the Ramachandran plot according to various secondary structures as well as some special angles for amino acids.....	65
<u>Figure 4.8</u> Ramachandran Angles for (A) 1,(B) 10,(C) 100 replica series.	66

Chapter 1: Introduction to vitamin D₃ & milk proteins

1.1 Focus of the thesis

Vitamin D₃ (VD₃) allows absorption of calcium and phosphate ions in the intestine by promoting production of ion transport proteins¹. Although milk is a great source of calcium and phosphate, it is deficient in vitamin D₃ (~1.56nM)^{2,3}. Without vitamin D, the hydroxyapatite ions (Ca²⁺ and PO₄³⁻) are not absorbed at low concentrations⁴. In response to this, many countries including Canada, mandate vitamin D fortification in milk⁵. The fate of vitamin D after fortification – whether it binds to caseins and to whey proteins or remains bound to the carrier molecule – is unclear. This information is important for reliable transfer of vitamin D from fluid milk to popular dairy products such as cheese⁶⁻⁸. Since milk consumption is decreasing, expanding the dietary sources of vitamin D is a public health issue as only 4% of Canadians have adequate levels of vitamin D⁹. Cheese is a good source of calcium, contains only traces of lactose and has a high demand amongst consumers¹⁰. Consumption of cheese has been steadily rising and is projected to continue to rise, owing to its variety (Figure 2.2b)¹⁰. A simple solid content calculation can show that in a year, an average Canadian consumes about twice the solids from cheese rather than from milk (circa 12kg vs. 6kg, respectively)¹⁰. Given that cheese is more nutritious on a per gram basis and its consumption is increasing, cheese could potentially be the next food to be mandated for fortification.

Unfortunately, transfer of VD₃ from fortified milk to cheese is unreliable and ranges from 39 - 90% retention^{6,7}. There are three reasons to explain this variability. First, whey

proteins in milk can bind hydrophobic ligands including VD₃^{11, 12}. So, during cheese pressing when whey is removed, there is a consequent loss of serum whey proteins and any bound VD₃. Second, milk fat plays a significant role in retention of VD₃ with a very predictable correlation – the higher the fat content, the higher the VD₃ retention – as the two can associate via hydrophobic interactions⁷. Finally, the current VD₃ fortification method in milk uses a microemulsion stabilized by polysorbate 80 (P80). This method is only effective in aqueous solutions, and therefore may be a cause of VD₃ loss during cheese making. At present, the most effective method of incorporating VD₃ into cheese is to utilize the casein fraction of milk which is also responsible for the structure of cheese¹³. Although this potential has been shown previously¹⁴, it has not been compared to whey proteins, or to P80 which would allow the determination of best fortification practices. This thesis shows that casein can bind to vitamin D₃ and if optimized in the future, the technology could potentially facilitate reliable fortification of dairy products.

Caseins, which are naturally found in the form of micelles, are known as nature's nanoparticles as they are designed to deliver calcium phosphate in milk¹⁵. As micelles, they can be described as supramolecular protein structures consisting of α_{s1} -casein, α_{s2} -casein, β -casein, and κ -casein at a ratio of 4:1:3.5:1.5 by molarity¹⁶. The internal structure of each micelle is maintained by hydrophobic interactions between protein chains and *via* electrostatic bridging through calcium phosphate¹⁵. The micelles maintain colloidal stability in milk serum by the help of κ -casein, which acts as a polyelectrolyte brush providing intermicellar steric repulsion¹⁷.

Structurally, caseins are classified as intrinsically disordered proteins (IDPs)¹⁸. The location of each casein protein on the spectrum of structural disorder is unknown, but the

degree of disorder dictates their function¹⁸. In the mammary cell, micelles are initially formed by weak van der Waals interactions while trapped within membrane-bound vesicles¹⁹. These vesicles are secreted from the Golgi apparatus and travel to the apical membrane of the mammary cells where they are released into the extracellular matrix¹⁹. During this transit, the vesicles pass through the cytoplasm where they encounter high concentration of calcium and phosphate ions²⁰. The intrinsic disorder of such weakly-formed micelles allows fast sequestration and binding of the ions, which leads to the formation of amorphous calcium phosphate nanoclusters^{18, 21, 22}. This prevents ectopic calcification of the mammary glands and allows active transport of the insoluble calcium phosphate salt necessary for skeletal development of neonates²³.

Disordered proteins are abundant in nature which underscores the need for detailed structure function studies. For caseins, in addition to localization of ligand binding, knowing the structure can lead to an understanding of internal organization of the micelle. However, to date, there is disagreement even with structural studies coming from the same lab using different equipment. This stems from the fact that many structure determination techniques depend on databases of globular proteins for structural assignments. For IDPs, use of these databases skew the final conformation calculation towards folded, globular structures and away from relatively unstable conformations such as the polyproline-II (PPII) helix and extended β -strands²⁴. This is the case for circular dichroism and FTIR. Nuclear magnetic resonance (NMR) has also been used to characterize structure of disordered proteins, but they have to be small, and manual characterization of chemical shifts is required to determine the structure.

Molecular dynamics (MD) simulation has been hailed as the ‘holy grail’ of protein structure determination²⁵. IDPs are difficult to study structurally due to conformational heterogeneity and their tendency to aggregate²⁶. Computer simulations, which are not limited by these experimental constraints, provide atomic resolution in studying IDPs. In testing out MD simulations for casein structure determination, it is important to validate the method with experimental results. Thus far, MD has been successful in characterizing the folding pathways of small polypeptides. So to perform MD on casein proteins, one has to start small and validate results with the most well studied casein polypeptide. It has been shown that β -casein phosphopeptide (f1-25) (β -CPP) is an appropriate structural and functional analog of the peptide chains surrounding the calcium phosphate nanoclusters in the casein micelle²⁷. Qualitative structural changes in the peptide have been observed using NMR upon binding of calcium phosphate ions²⁸. Conventional structural studies of the peptide to date, however, have yielded no agreement on the secondary structure. Hence, β -casein phosphopeptide (f1-25) was a good starting point for MD simulations with caseins. As part of this thesis, MD simulation of β -CPP accomplished three major goals. First, it allowed validation of the MD simulation method in comparison to experimental results. Second, it helped resolving the secondary structure controversy of β -CPP that currently exists in literature. Finally, with further simulations it will assist structure function correlations of the peptide in its biological role as a calcium carrier.

1.2 History of the thesis

This thesis started out with the goal of optimizing the binding interaction of casein micelles with vitamin D₃ (VD₃). This was to be accomplished by studying their structure-function relationship (Figure 1.1) and to localize the casein micelle moieties involved with VD₃ binding. Experimentally, the interaction can be understood by characterizing thermodynamics of binding. This required that the association constant (K_d) and stoichiometry of binding (n) be measured. As fluorescence had been used extensively in literature, a significant amount of time and

effort was initially spent using this technique to study VD₃-casein interaction.

However, the end results proved unreliable, due to technical caveats. These included issues with data acquisition (convolution of different mechanisms:

inner filter effect, structural changes, and energy transfer), processing and analysis

(multiple ways to process data, and issues with equation derivations). After unsuccessful attempts at using fluorescence, isothermal titration calorimetry (ITC) was explored. This method avoids many of the caveats of fluorescence and is considered the gold standard in studying binding interactions. The main requirement of this method is that of an appropriate solvent for both VD₃ and casein, so that the response observed is only due to the interaction. Although many solvents and combinations were tested, ultimately, none were suitable for this system. This was because VD₃ is hydrophobic whereas caseins are

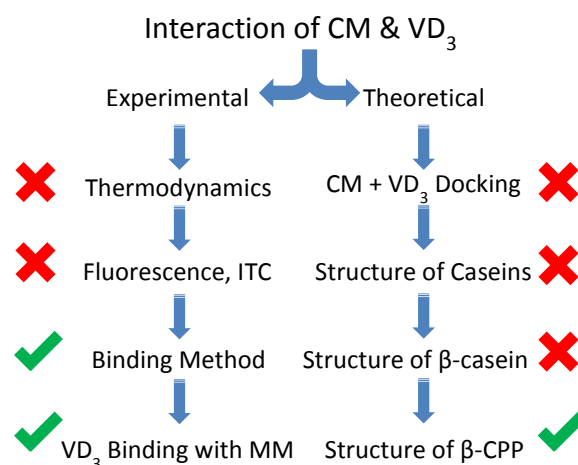


Figure 1.1 History of M.Sc. Thesis. CM: Casein Micelle, VD₃: Vitamin D₃, ITC: Isothermal Titration Calorimetry, MM: Macromolecules; β-CPP: β-casein phosphopeptide (f1-25).

primarily hydrophilic. So, a new binding method was developed to quantify the interaction of VD₃ with caseins. At that point, the focus of experiments became the legitimacy of the newly developed assay by testing it with various macromolecules known to bind vitamin D₃.

On the theoretical side, initial attempts with docking of VD₃ on caseins couldn't be accomplished, because of the lack of structures available for the complete micelle. For docking experiments, the primary prerequisite is to have a high quality structure native to the solvent conditions in which the experiment is to take place. As yet, only one set of *in vacuo* simulation has been conducted per casein²⁹⁻³¹. Simple explicit water simulations revealed that structures were drastically different from *in vacuo* work done in the past (data not shown). However, starting from different structures, the simulation results were also significantly different. This is a fundamental characteristic of disordered proteins in that these proteins have a multitude of conformations with similar energies. In order to describe the structure of disordered proteins, rather than specifying convergence to a singular structure, there is a need to describe the conformation ensemble³². This is computationally very expensive and time consuming. Knowing that the complete folding ensembles through MD have been established for small proteins only, it was decided that it would be wise to develop a method for molecular dynamics of the most experimentally well studied and functional part of caseins, the β -casein phosphopeptide (β -CPP). This way the simulations could be validation though experiment. Furthermore, β -CPP on its own is an industrially and medically important peptide²⁷, simulation of which would provide insight into its conformational preferences.

1.3 Aim of the thesis

Due to the challenges encountered, there were two goals of this thesis. The first was to demonstrate that vitamin D₃ binds to the casein micelle and to compare this binding affinity to other macromolecules in milk, thus establishing the possible fate of vitamin D in fortified dairy products. The second goal emerged from the dearth of knowledge on the internal organization of the casein micelle and, by corollary, that of individual caseins. To this end, MD simulations were used to determine the conformation ensemble of β -casein phosphopeptide (f1-25) in its native environment and to validate the method by cross-experimental correlation.

The thesis begins with a review of the structure and function of VD₃ and milk proteins, with special emphasis on caseins (Chapter 2). This is followed by the development of an assay for VD₃ binding to test its efficacy and limitations (Chapter 3) where the results of VD₃ binding in the presence of caseins and other food-based macromolecules are explored. Finally, to improve upon the current structural disagreements on casein structure, the structure of β -casein phosphopeptide (f1-25) using MD simulation is reported (Chapter 4).

Chapter 2: Structure & function of vitamin D₃ and milk proteins

2.1 Vitamin D₃

In the 1920s, it was recognized that amelioration of rickets status could be achieved by periodic exposure to sunlight or administration of cod liver oil^{2, 11, 33-35}. Within a short time of that realization, the principal antirachitic component was isolated and named vitamin D³⁶. The process of irradiation of patients as well as foods to provide antirachitic protection was patented in 1925 by the University of Wisconsin⁴. Shortly thereafter, a great deal of work dealing with chemistry of vitamin D ensued, which led to the assignment of vitamin D₂ and D₃ as 28 and 27 carbon steroids, respectively⁴. Chemical progenitors of these two compounds reflect their aetiology.

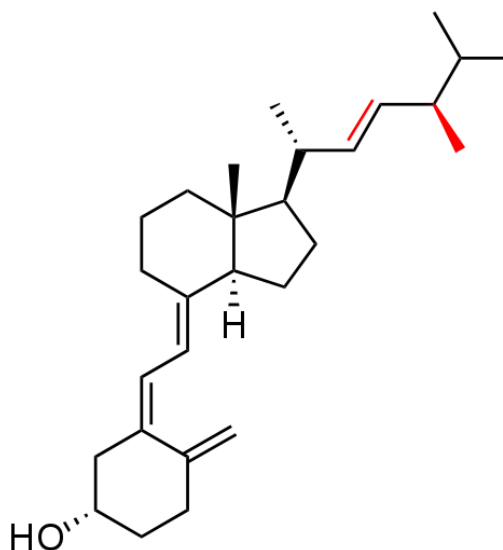


Figure 2.1 Chemical Structure of Vitamin D₂. Vitamin D₃ does not have the double bond and methyl group indicated in red.

Ergocalciferol is derived from ergosterol which occurs especially in fungi (hence the prefix ergo- from the group of fungi known as ergot) as well as in yeasts and plants³⁶. The etymology of cholecalciferol, on the other hand, comes from its precursor compound cholesterol which derives its prefix chole- from Greek meaning bile, pointing to the fact that it is anabolized in the liver and sent to the gall bladder to be stored as bile³⁷. So the term *vitamin D* is a catch-all for seco-steroids (D₂ – ergocalciferol, D₃ – cholecalciferol)

in which one of the steroid rings has been broken (*seco* is latin for *to cut*). Vitamin D₂ and D₃ form through UV-B (270-300nm) irradiation of their respective sterols; ergosterol in fungi, yeasts and plants and 7-dehydrocholesterol in humans^{3, 38}. The difference between the two vitamin D derivatives is only in the double bond and methyl group which are present in D₂ but absent in D₃ (Figure 2.1).

In humans, activation of VD₃ involves hydroxylation at the 25th carbon in the liver followed by a second hydroxylation at the 1st carbon in the kidney^{1, 4}. This activated compound is a hormone commonly referred to as 1,25(OH)₂D or calcitrol. Once released into circulation, calcitrol is transported by vitamin D-binding protein (VDBP) to various organs³³. The ultimate molecular target of calcitrol is the vitamin D receptor (VDR) found in the nuclei of cells¹. After calcitrol binding, VDR becomes activated as a transcription factor, modulating gene expression for specific absorption and transport proteins destined for the distal large intestine³⁹. These proteins transport calcium and phosphate ions from the intestine and into the bloodstream. The skeletal system incorporates excess ions through accretion in bones². Hence, VD₃ maintains homeostasis of blood calcium and phosphates through dietary absorption from the duodenum and jejunum⁴⁰. In the case where these dietary ions are unavailable, VD₃ can stimulate bone resorption to maintain homeostasis². Aside from this classic role, VD₃ is involved in osteoblast formation, fetus development, pancreatic function, neural function, and immunity³⁴. Positive benefits of prolonged high serum VD₃ levels are associated with lower rates of malignant cancers, and of multiple sclerosis^{3, 7, 38}.

2.1.1 Nutrition

In November 2010, an Institute of Medicine (IOM) committee commissioned jointly by the governments of United States and Canada, published a report regarding population needs of vitamin D. The committee conducted a comprehensive review of evidence regarding benefits of VD_3 to skeletal outcomes⁴¹. Although benefits of VD_3 are pleiotropic³⁴, for methodological uniformity the committee focused on bone health as the indicator of physiological outcome. Based on the review, the committee concluded that recommended dietary allowance (RDA) per day for VD_3 should be 400 International Units (IU) for persons between the age of 0-12 months, 600 IU per day for ages 1-70 years and 800 IU/day for individuals older than 70. Clinically, VD_3 status of an individual is based on serum blood levels of $25(\text{OH})\text{D}$ ⁸. These RDAs correspond to serum $25(\text{OH})\text{D}$ levels of 50nM. For reference, 1g (386.7g/mol) of VD_3 is 40 million IU.

The revised RDAs were a modest increase from the past, but immediately after the review was published, many scientific experts and organizations criticized its outcomes. This emerged from the confusion where certain studies pointed out the benefits of high serum $25(\text{OH})\text{D}$ whereas others showed benefit at low levels of $25(\text{OH})\text{D}$. Another point of contention was the claim that the revised RDA recommendations were good for 97.5% of the population, with the remaining 2.5% high risk individuals with higher VD_3 requirements to have sufficient $25(\text{OH})\text{D}$ blood levels. This list included individuals suffering from malabsorption, celiac sprue disease, inflammatory distal bowel disease, osteoporosis, obesity, *etc.*⁴². As well, the IOM report suggested that the North American population mean for serum $25(\text{OH})\text{D}$ was ~50nM, concluding that there was no need for supplementation. However, if the population mean for serum $25(\text{OH})\text{D}$ was 50nM, and the adequate cut-off for sufficient $25(\text{OH})\text{D}$ is also 50nM, this would suggest that half of

the population was deficient. Many leading experts have suggested that the optimal 25(OH)D blood levels should be ~75nM to mitigate risks of fractures⁴². Based on this estimate, the 2010 Statscan study shows that during the period of 2007-2009, only 4% of Canadians are meeting their vitamin D needs⁹.

2.1.2 Barriers to adequate levels of VD₃

Multiple barriers prevent adequate VD₃ status in the Canadian population⁵. The major source of VD₃ is the sun, exposure to which initiates cutaneous synthesis of VD₃¹. This major source is hindered by geographical, environmental, cultural, and physiological factors^{2, 5}. There are very few dietary sources of VD₃, with many sources not regularly consumed (*e.g.*, cod liver oil, fatty fish)³⁸. To compensate for the lack of sunlight and paucity of VD₃ in nutritional sources, reliance on fortified foods and supplements has increased⁵. Fortification of VD₃ is mandatory in Canada for fluid milk (100IU/250mL) and margarine (53IU/10g)⁵. Cross-sectional studies suggest current fortification practices are ineffective given antiquated fortification guidelines and under-fortification of regulated foods⁴³. As well, fluid milk consumption has been steadily declining over the past decades and forecasts indicate regressing demands (Figure 2.2a)¹⁰. Lactose intolerance and maldigestion has become increasingly commonplace in Canadians⁴⁴⁻⁴⁶. Scientific evidence and population needs thus necessitate development of additional foods fortified with VD₃ as cost-effective measures to improving public health.

2.1.3 Sources

Cutaneous exposure of skin is the major source of VD₃ for humans³⁹. A variety of geographical, environmental, cultural, and physiological factors determine the capacity to produce VD₃ which include extreme latitudes, seasonal changes, increased melanin pigmentation, topical application of sunscreen, or cultural influences on dress code⁵.

Ageing also reduces VD_3 production capacity due to decreased precursor concentration. Of the natural sources of VD_3 , cod liver oil has the highest concentration, followed by fatty fish, mammalian liver, eggs and dairy products⁴⁷. Although the benefits of cod liver oil have been known for generations, it is not normally consumed³. As for other sources, VD_3 is only available in trace amounts⁴⁷. Hence natural sources are not viable for recommended dietary VD_3 intake³⁸. The third source is supplementation of VD_3 in diet either through vitamin D capsules or through food fortification⁸. These two have become increasingly popular in recent years given reduced exposure to sunlight in North America⁴⁸.

2.1.4 Fortification of dairy products

Milk intended for processed dairy goods is not fortified with VD_3 ⁵. In Canada, regulations regarding fortified products are overseen by the Canadian Food Inspection Agency (CFIA). To increase public health, the CFIA has mandated the fortification of fluid milk with vitamin D at 100IU/250mL or 25nM⁵. Unfortunately, evidence indicates that milk fortification alone is inadequate due to out-dated fortification guidelines as well

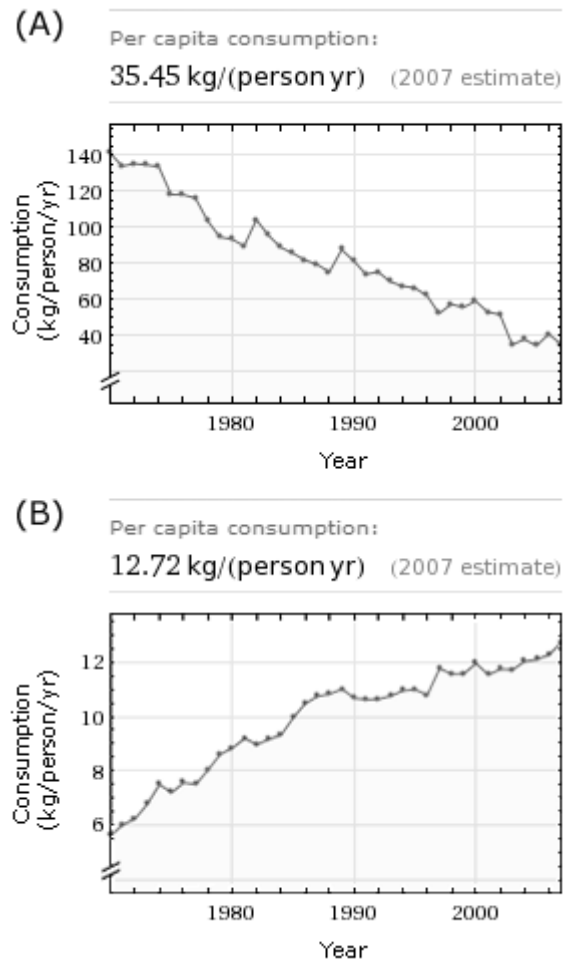


Figure 2.2 History of (A) milk and (B) cheese consumption in Canada. Adapted from Wolfram Alpha.¹⁰

as under-fortification^{5, 43}. As mentioned, fluid milk consumption has steadily declined from 140kg/person/year in 1970 to 35kg/person/year in 2007¹⁰ (Figure 2.2). Furthermore, variable VD₃ levels have been found in fortified milk, with many below the stated value. A study of Ontario retail milk samples conducted in 2000 found that only 20% of skim milk, 40% of 2% milk fat, and 20% of whole milk contained the recommended levels⁴³. It is unknown whether vitamin D variability is due to inconsistencies in fortification or to instability of the vitamin itself.⁴³

Conversely, cheese consumption has steadily risen over the years, most likely owing to variety and its versatility of use in foods¹⁰ (Figure 2.2). Cheese is a good source of calcium and contains only traces of lactose. However, transfer of VD₃ from fortified milk to cheese has proven unreliable⁶⁻⁸. The current fortification methods involve addition of an emulsion containing VD₃ solubilized by P80 to the milk. Comparative analysis of three fortification methods for Cheddar cheese demonstrated 39-59% of the initial VD₃ load was lost, depending on fortification method, with loss attributed to processing conditions that could lead to oxidation⁶. Hence, the speculation is that fortification using the current emulsification technique cannot sufficiently protect VD₃ during processing. It may be possible to protect VD₃ by binding it to other food-grade carrier molecules such as caseins or whey proteins^{12, 14}. The fate of VD₃ in fortified milk, whether it becomes involved in competitive binding with caseins and whey proteins, or whether it stays bound to the carrier molecule (P80), remains unclear.

2.2 Milk proteins

Milk proteins can be broadly classified into casein and whey proteins. In bovine milk, the dry protein weight fraction consists of 20% whey and 80% casein. Whey protein is a mixture of many small, globular proteins whereas the casein (cn) fraction is a 4:1:3.5:1.5 weight ratios of α_{s1} , α_{s2} , β , and κ -casein.⁴⁹

2.3 Whey proteins

Whey is generated as a by-product during cheese processing (in the pressing step) and contains a mixture of globular proteins that can be isolated by membrane filtration followed by spray drying⁵⁰. The protein isolate, usually known as whey protein isolate (WPI) has four major fractions: β -lactoglobulin (β -Lg) (65%), α -lactalbumin (25%), serum albumin (8%) and low concentrations of immunoglobulins (2%)³⁵. Of these major fractions, only β -lactoglobulin has shown propensity towards VD_3 binding^{11, 12, 51, 52}.

2.3.1 Structure of β -lactoglobulin

Bovine β -Lg consists of 162 amino acid residues with a molecular weight of ~18 kDa. Due to sequence and structural homology, it belongs to the *lipocalin* family, which can bind hydrophobic ligands.⁵³

The structure of β -Lg is composed of eight up and down antiparallel β -strands folded into a flattened cone or a calyx (Figure 2.3). The β -sheet barrel is surrounded by three α -helices, one at the bottom of the calyx and two at the top. There are seven connective loops three of which are short and rigid while the remaining four are long, flexible and mobile. The calyx is closed off by one α -helix (Trp-19) while the mobile loop (residue 85-90) acts as an access 'lid' to the hydrophobic barrel. The stability of the tertiary

structure is governed by two disulphide bonds formed by four out of the five cysteine residues.⁵⁴

2.3.2 Function of β -lactoglobulin

It has been shown that β -Lg is responsible for transport of small hydrophobic ligands which, together with the evidence of its fast catabolism³⁵ stability at low pH and binding at high pH⁵⁵, has led to the hypothesis that the evolutionary function of β -Lg is that of ligand transport^{11, 12, 54}.

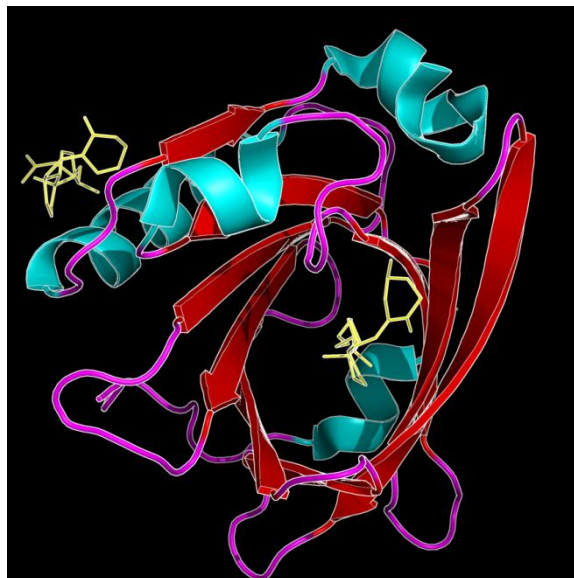


Figure 2.3 Crystal Structure of β -lactoglobulin adapted from the RSCB database, showing the two binding sites for VD_3 (yellow), one in the red β -sheet barrel and the other on the surface of the hydrophobic barrel and in line with the adjacent α -helix.¹¹

The structure of β -Lg is affected by changes in pH and temperature. The changes in pH are associated with concomitant electrostatic changes associated with the protonation state of negative and positive amino acid residues⁵⁵. These affect protein denaturation (and coagulation), dimerization and ligand binding properties^{50, 54}. The monomer-dimer equilibrium of β -Lg is a direct function of pH in which the range for dimerization is from 3.0-8.0, outside which it becomes monomeric. In the dimeric state, the units are linked *via* twelve hydrogen bonds. The most important pH-induced conformational change is the Tanford transition occurring in the loop forming region (residues 85–90) between pH 7.5 - 8.5⁵⁶. This conformational change, which is brought about by protonation of Glu89, mediates access to the central calyx and allows entry of hydrophobic ligands into the β -sheet barrel⁵⁵.

2.3.3 Ligand binding

β -Lg can bind hydrophobic ligands at two sites – one inside the central calyx and the other in a groove formed by an adjacent α -helix and β -barrel. Crystal structures of β -Lg with bound retinol, retinoic acid, palmitic acid and vitamin D₃ have been discovered. All related studies have been conducted at pH 8.0 at which the hydrophobic calyx of β -Lg is accessible.^{12, 51, 52, 55, 57-59}

2.4 Caseins

The name *Casein* comes from the latin word *caseus* which means cheese, pointing to one of the end products of these proteins in the body of ruminant mammals. The four main casein proteins are named according to their electrophoretic mobility.¹⁶

2.4.1 Structure

The most important function of casein is the transport of the insoluble calcium phosphate ($K_{sp} 2.70 \times 10^{-7}$). This is accomplished by the formation of soluble micelles that sequester the calcium through phosphoserine groups on caseins. The internal structure of the micelle is highly-debated⁵⁰. The inherent structural heterogeneity of caseins, imparts a characterization challenge that is relatively more demanding than for a globular protein.

The different amino acids are colour coded to five distinct types according to the key in Figure 2.4**Error! Reference source not found..** Caseins are mostly uncharged proteins with charged moieties dotting the chain. The proteins, α_{s1} -cn and α_{s2} -cn contain the highest number of charged sites and the highest number of phosphorylated centres (PCs); β -cn contains a high proportion of neutral residues while κ -cn contains the neutral residues towards its N-terminus region prior to the chymosin-sensitive bond (105/106).

2.4.2 *Function*

Casein micelles have multiple functions which are related to their origin (mammary secretory cells) and destination (young, since it is their only mode of nutrition acquisition). Once ingested in the milk solution, the caseins forms a gel as the pH of stomach is less than the isoelectric point of caseins (pH 1.5 vs. pI 4.6). Gel formation occurs due to neutralization of electrostatic charges near

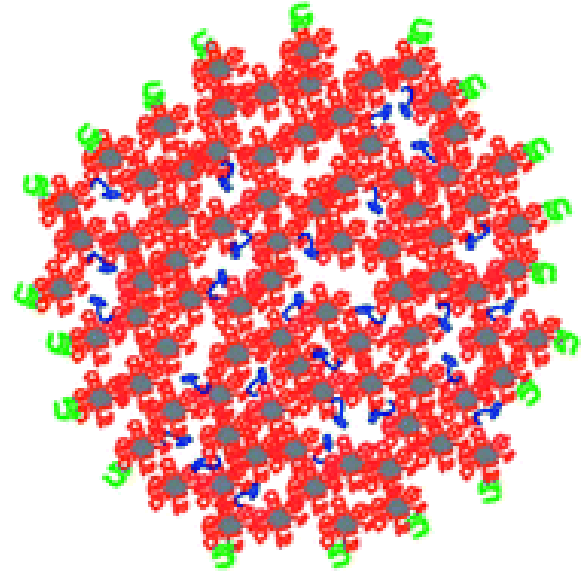


Figure 2.5 Hypothesized model of the casein micelle with calcium phosphate nanoclusters (grey), α s1, and α s2 casein shown in red, β -casein shown in blue, and κ -casein shown on the surface in green. Reprinted with permission from Royal Society of Chemistry.⁴⁹

its pI. For ruminant mammals, the gel is formed in the stomach through cleavage of the κ -cn polyelectrolyte brush by chymosin. This results in the loss of intermicellar steric repulsion, and more collisions of the core micelle. At that core level distance, van der Waals interactions become significant and the gel formation is initiated. Even though these gel formation mechanisms differ, the outcome is the same. Gel formation in the stomach slows down breakdown and increases calcium bioavailability by increasing gel residence time in the GI tract. As the gels degrade, they release many essential amino acids^{50, 68}. Caseins contain all essential amino acids and their proteolytic digestion ensures a rich supply of amino acids to the host. In addition, since the gel is degraded through piece-wise peptide cleavage, bioactive peptides are also formed directly in the stomach.

The assumption that a protein requires a set 3D structure to function has been superceded by rising evidence showing that many disordered proteins have specific functions⁶⁹.

These functions can be categorized into five broad classes: entropic chains, effectors, assemblers, display sites and scavengers⁷⁰. In this scheme, the caseins are classed as scavengers of calcium phosphate. To avoid confusion with

Table 2.1 Structure Function correlation of the Casein Micelle⁷³

<i>Function</i>	<i>Structure</i>
Amino acid supply	Protein digestion
Viscosity reduction	Surface activity
Acid Gel	Coagulation at pH 4.6 ¹³
Rennet Gel	Cleavage of κ -cn ⁷⁵
Calcium transport	P-SER and P-THR
Calcium bioavailability	GI tract residence time
Bioactive peptides	Cleaved caseins
CaHPO ₄ sequestration	Intrinsic disorder ⁶⁰

denatured proteins, caseins are described as *rheomorphic* proteins⁶⁰. The paucity of secondary structure is due to the lack of cysteine residues for disulphide bridging and an abundance of proline residues⁷¹. Hence, casein conformation is essentially dictated by the local environment where their flexible nature allows them to ‘go with the flow’⁷². The rheomorphic nature of caseins allows fast sequestration of calcium species by providing nucleation sites at phosphorylated serines or threonines, thereby preventing pathological calcification of secretory cells⁷³. More generally, such conformational flexibility permits stable complexation with calcium phosphate much more rapidly than a globular conformation could allow⁷⁴.

There are three amino acids in eukaryotes that are phosphorylated – phosphoserine, phosphothreonine and phosphotyrosine⁷⁶. Of these, phosphoserine is the most common, followed by phosphothreonine and the rare phosphotyrosine. Complexation of calcium occurs through phosphorylated sites on caseins. The calcium binding capacity of the various caseins is proportional to their molar phosphate content. This type of behaviour is synonymous with structural proteins that form templates for mineralized hard shells, eg. calcium-binding proteins for bone (acidic glycoprotein-75, osteopontin, and bone sialoprotein)⁷⁷, dentin for teeth⁷⁸, exoskeletons of certain organisms⁷⁹.

Industrial exploitations of casein are evident in the myriad variety of dairy products available on the shelves of grocery markets (*e.g.*, cheese, yogurt, and ice-cream). Current trends in micellar engineering include utilization of casein micelles as nanoscopic vehicles for nutraceutical and drug delivery^{14, 80} due to their generally-regarded-as-safe (GRAS) standard, nanoscopic size⁶⁴, UV-oxidation protection¹⁴ and controllable self-assembly^{81, 82}.

Of all caseins, the phosphorylated peptide of β -casein f(1-25) has recently come under increased scrutiny due to its ease of isolation and stability during processing²⁷. Functionally, this peptide can be used for calcium phosphate fortification or medical applications such as teeth biomineralization^{83, 84}. There is a lack of consensus over its structure, owing to its membership to the class of intrinsically disordered peptides (IDPs). A key reason why the structure of this peptide is important is because it contains a major epitope for IgE binding associated with milk allergy. With 2.5% of infants affected by this allergy, knowledge of its structure becomes medically-relevant^{85, 86}. Compilation of data over the past thirty years suggests minimal consensus on structure (see Table 4.1) due to the difficulty in experimentally studying IDPs as they are structurally heterogeneous²⁴. Computer simulations have recently emerged as a useful tool to understand disordered peptides as they are not limited by experimental constraints that result in structural heterogeneity and aggregation²⁶.

Chapter 3: Binding of vitamin D₃ to food-based macromolecules

3.1 Abstract

Most strategies for fortification of dairy products with vitamin D₃ (VD₃) are based on emulsification. However, studies have shown lack of VD₃ retention during Cheddar cheese processing and possible oxidation during storage using this method. There is thus a need to develop a more robust fortification strategy. As a first step towards achieving this goal, we present the development of a simple approach to characterize the binding of VD₃ to food-related macromolecules (MMs) that may be used for enrichment. We evaluated VD₃ binding to sodium caseinate (SCN), hydroxypropyl- β -cyclodextrin (HP β CD), whey protein isolate (WPI) and polysorbate 80 (P80). At 10g/L MM, the binding affinity was P80>SCN>HP β CD>WPI. However, at concentrations \leq 1g/L, SCN provided the most effective binding to VD₃ as it was below its critical micelle concentration (CMC), where its solvent-accessible hydrophobic surface area was highest. Above its CMC, SCN showed preferential binding to other caseins rather than VD₃. WPI did not show any significant binding. HP β CD, on the other hand showed excellent binding (0.053mg/mL at 100g/L). These results suggested that strategies using SCN and/or HP β CD could provide an alternative method for fortifying cheese with VD₃. In addition, as their protective abilities are known in the literature, these two MMs may also protect VD₃ during processing of Cheddar cheese.

3.2 Introduction

Vitamin D₃ (VD₃) deficiency is a largely unacknowledged public health issue, both in Canada and abroad⁴⁰. Chronic VD₃ deficiency prevents proper bone mineralization, growth and remodelling and can lead to rickets in children and osteoporosis or osteomalacia in adults³. Clinically, VD₃ status is measured by 25-hydroxyvitamin D – 25(OH)D. Low levels of 25(OH)D (<50nM) are associated with increased risk of incidence and mortality from various cancers while high levels (≥75nM) are associated with inverse incident risks of rheumatoid arthritis, microbial infections, cardiovascular disease, diabetes and multiple sclerosis³⁴. Yet, meta-analysis of VD₃ correlations with health and disease studies has revealed that ≥75nM of 25(OH)D concentration in blood is necessary to reap the most health benefits from VD₃ consumption^{3, 8, 9, 38, 40, 87}. Yet, only 4% of the Canadian population is above this threshold⁹.

3.2.1 Necessity for an alternative assay to fluorescence

The development of new and improved VD₃ fortification technologies based on dairy proteins requires that an appropriate method of evaluating binding affinity be used. In the recent past, fluorescence and isothermal titration calorimetry have been evaluated as tools to determine protein-ligand interactions. However, these standard methods require a consistent solvent composition to avoid complicated corrections for subsequent data analysis and to ensure that changes in signal response truly arise from protein-ligand interactions. This is especially difficult if the ligand is hydrophobic and protein hydrophilic.

One of the techniques used to determine the thermodynamics of the casein-VD₃ interaction is based on the changes in protein-Tryptophan fluorescence following the

addition of VD₃ dissolved in ethanol to proteins in aqueous solution. Even though a decrease in fluorescence is observed, the decrease may not be due to the binding interaction, but rather from a reduction in excitation and emission intensity due to scattering⁸⁸. This is known as the ‘inner filter effect’. If a hydrophobic substance dissolved in an organic solvent is titrated into an aqueous solution with protein, the result is a colloidal suspension that scatters light at both excitation and emission wavelengths. The term originates from this phenomenon – *inner* indicating that the effect is inherent in the sample and *filter* refers to the reduction in excitation and emission light intensity as it enters and exits the sample⁸⁸. This effect can be avoided by reducing the concentration of the titrant, (eg. VD₃) but this invalidates equations used to model changes in response⁸⁸. As this is not the topic of this thesis, a good reference on this matter is written by van de Weert, 2010⁸⁸.

Another technique that is useful in studying protein-ligand interactions is isothermal titration calorimetry (ITC). This technique is considered the gold standard in studying thermodynamics of interactions⁸⁹⁻⁹¹. The prerequisite for ITC is that the solvent for the ligand and for the protein be identical to avoid enthalpy changes arising from the use of dissimilar solvents (solvation energy)^{90, 92}.

Given the inherent requirement of identical solvent conditions for both techniques, the study of VD₃ binding to proteins using these techniques is challenging, as this vitamin is very hydrophobic while milk proteins are generally hydrophilic. The efficient assay herein developed is an indirect measurement of the interaction between VD₃ and various macromolecules that avoids the complication of having dissimilar solvent conditions.

The basic premise of this approach is based on the shaken-flask and solvent evaporation methods used in the pharmaceutical industry⁹³. In short, a known quantity of VD₃ is added to a test tube and shaken with a solution containing the macromolecule of interest. The macromolecule solution is removed and the undissolved VD₃ is quantified.

3.2.2 Selection of test molecules

The macromolecules tested with the developed assay were selected based on VD₃ fortification in raw milk and its lack of retention in dairy products. The common fortification method (with a polysorbate 80-based microemulsion) is unreliable for cheese manufacture as the added VD₃ is not protected against acidification and fermentation conditions^{6, 94}. The goal of this study was to determine whether other possible food macromolecules could be used as VD₃ binding agents at the native pH of milk (pH 6.7). These macromolecules included dairy proteins (sodium caseinate, whey protein isolate), and hydroxypropyl- β -cyclodextrin. Binding of VD₃ with P80 was also determined.

Sodium caseinate (SCN) contains four proteins that arrange into micellar structures when hydrated⁴⁹. The micellar structure is capable of binding and transporting calcium and is also the main ingredient that provides the structure and texture of processed dairy foods – yogurt, cheese, and ice cream^{6, 7, 13, 17, 51}. Moreover, casein has been shown to bind with vitamin D₂¹⁴ and D₃⁵¹ and has shown to protect vitamin D₂ against UV degradation¹⁴. Whey protein isolate (WPI) consists of the proteins that remain in the milk serum after the caseins are removed. WPI primarily consists of α -lactalbumin and β -lactoglobulin, with the latter implicated in binding and transport of several hydrophobic ligands, including VD₃^{11, 12, 51, 52, 57, 95, 96}. Denatured whey protein isolate (dWPI) solutions, obtained by heating at 70°C for 1hr, were also tested as denaturation of β -lactoglobulin

has been shown to occur during pasteurization of milk^{97, 98} which may alter VD₃ binding equilibrium between caseins and whey proteins. Polysorbate 80 (P80) is a food-based emulsifier that forms the basis of the currently used method to fortify raw milk with VD₃. It has a very high emulsifying capacity even at low concentrations. It is able to bind to high concentration of hydrophobic molecules because of its low CMC (0.012mM). As it has the ability to form micellar structures, it offers a good comparison to casein micelles in terms of binding mechanism. Finally, hydroxypropyl- β -Cyclodextrin (HP β CD) is a molecule similar to β -lactoglobulin in that both have an internal hydrophobic binding site⁵⁸ and both can bind VD₃⁹⁹.

3.3 Materials and methods

VD₃ (>99%), ethanol (EtOH) (>95%), SCN, P80, HP- β -CD, NaCl, NaH₂PO₄, and Na₂HPO₄ were obtained from Sigma-Aldrich (Oakville, ON, Canada) and used without further purification. WPI (92% protein content) was a gift from Fonterra Inc. (Mississauga, ON, Canada). Macromolecule solutions were prepared in 0.1M PBS buffer pH 6.7 (80mM NaCl, 5.65mM Na₂HPO₄, 3.05mM NaH₂PO₄). Figure 3.1 outlines the binding assay schematic. Initially, 0.1mL of a 10g/L stock solution of VD₃ in EtOH was dispensed into a 15mL Falcon tube (BD FalconTM, Fisher Scientific, Markham, ON, Canada), placed under vacuum (28-29 inHg) for 1hr (Step 1), and air-dried in a fume hood for 24hrs, leaving behind a layer of adsorbed VD₃ (Step 2). This initial adsorbed layer, which nominally contained 1mg of VD, is termed VDi for *initial* VD mass which is confirmed by EtOH control during the experiment. A 10mL MM solution is added to the primed tube (Step 3), which was then placed on a VWR rocking platform (Mississauga, ON, Canada). To ensure equilibrium conditions the platform was automated to rock for 24 hrs at a 15° tilt and a rocking rate of 20 cycles/min (Step 4). The time for equilibrium was determined using kinetic data at 5, 24, and 48 hrs. The solution added to the VD₃-primed tube was either a control – phosphate buffer (PBS), reverse osmosis (RO) water or EtOH – or a MM solution: SCN, P80, WPI, or HP β CD. The mass of VD₃ that was solubilized in the MM solution was termed VDs. After equilibration, the MM solution was decanted (Step 5), and the tube was washed 5x with 5mL of RO water to remove any traces of MM (Step 6). Washing was performed gently to prevent shear removal of VD₃ layer. This was done by tilting the falcon tube allowing the water to flow down test tube wall. After decanting the water from step 6, the residual water was evaporated for 24

hours to prevent VD₃ loss. The remaining VD₃ at the bottom of the tube was then resolubilized in 10mL EtOH (Step 7). The VD₃ concentration was calculated by measuring the absorbance at 240nm (Step 8), which has the largest dynamic range and provided good sensitivity (Figure 3.2). This *final concentration* was converted to mass (by multiplying by EtOH volume 10mL) and is termed VD_f. The relationship between the three variables is:

$$VD_s = VD_i - VD_f \quad \text{Eq 3.1}$$

With this equation, the mass of VD₃ in the macromolecule solution compared to the PBS control was evaluated and used as an indicator of binding. All tests were conducted at standard laboratory conditions (298K, 1 atm).

Hydrodynamics of P80, SCN and VD₃ in PBS were measured for particle size using the Malvern Mastersizer (Worcestershire, UK) and ζ -potential was measured using Brookhaven 90Plus Particle Size Analyzer (Holtsville, USA).

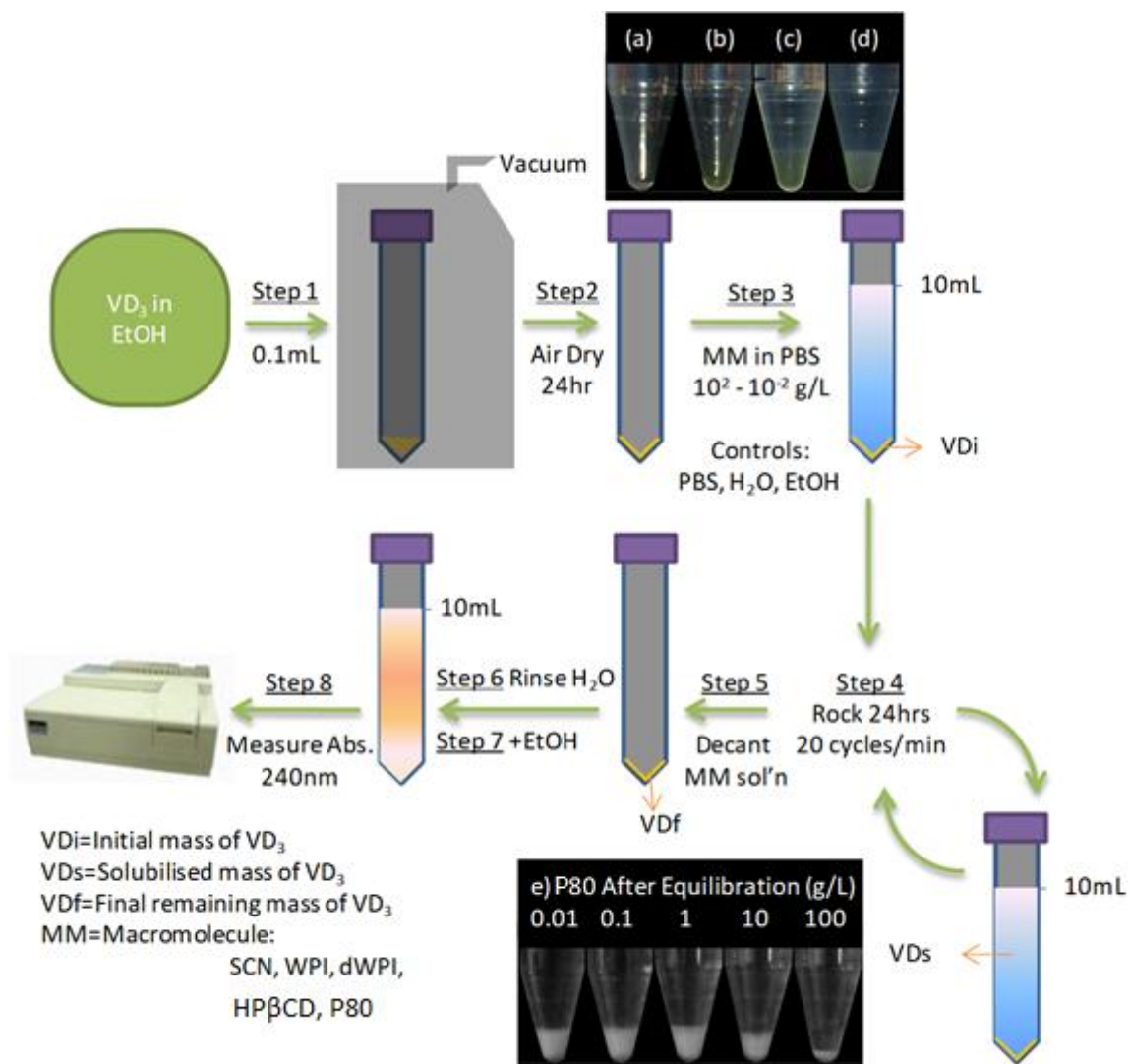


Figure 3.1 Vitamin D₃ Solubility Assay with step-wise pictures from P80 samples and 10mg of VD₃ layer. (a) Empty Falcon tube without VD₃ priming; (b) Ethanol dried VD₃-primed tube (c) Macromolecule solution added to VD₃-primed tube; (d) After a few minutes of addition of macromolecule solution; (e) After 24 hrs equilibration with P80 in increasing concentration of macromolecule from left to right.

3.4 Observations

3.4.1 Experimental remarks

When EtOH was evaporated from the test tube, a smooth, transparent light-yellow layer of VD₃ was formed (Figure 3.1b). This was desired, as layer-breakage was unwanted during the course of the experiment. As well, since VD₃ is hydrophobic, the test tube selection was based on a hydrophobic and inert plastic material. Tests for adsorption of

VD₃ with borosilicate glass test tubes showed that the VD₃ layer readily desorbed during experimentation due to the polarity of the substrate. Hence, plastic Falcon tubes were used for the experiments. Adsorption at the bottom of the Falcon tubes during the experiments (Figure 3.1, Step 4) was necessary to avoid the introduction of a filtration step and to minimize errors.

Upon adding the MM solution to the VD₃-primed tubes, the VD₃ layer started to become slightly white, and over 24hrs became completely white/cream-coloured (Figure 3.1c, d, e). Layer-whitening occurred due to exchange of VD₃ between the solid and the aqueous solution that “etches” the surface due to the movement of VD₃ molecules into the solution, thereby scattering the light crossing through the previously transparent layer and making it appear white and opaque. This phenomenon was independent of the type of aqueous solution added. Depending on the capacity of the MM to bind VD₃, layer thickness is also reduced (Figure 3.1e). Following equilibration, each tube was visually inspected for layer-breakage, for which there was no evidence. To further investigate whether there was microscopic layer breakage, post-binding recovery from the aqueous solution was performed (Section 3.4.3). Solubility was expressed in mg and was converted from mg/mL by multiplying all concentrations by 10mL.

3.4.2 Assay limits

Figure 3.2 shows the standard curve of VD₃ in EtOH measured at 240nm. This wavelength avoids absorbance from phenylalanine (257nm), tyrosine (274nm) and tryptophan (280nm)¹⁰⁰ and also allows the largest dynamic range for the measurement of concentration.

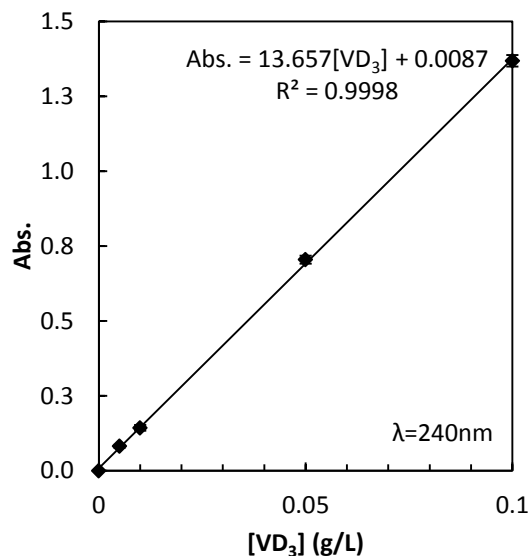


Figure 3.2 Standard Curve of VD₃ in EtOH

Furthermore, since all measurements of VD₃ were conducted in EtOH, it is unlikely that there would be any absorbance from the protein moieties, as the protein MMs were insoluble in EtOH. The large molar absorptivity coefficient, 5384.96 Lmol⁻¹cm⁻¹ results in a sensitivity of 13.6L/g. As a result of high absorbance, the dynamic range of the assay was stopped at 0.1g/L [VD₃]; further increases were non-linear. Although the absorbance range for conventional UV/Vis spectrophotometers is from 0 to 4 absorbance units (AU)¹⁰¹, deviations were observed at concentrations of 0.20 g/L which placed the absorbance response at 3 AU.

3.4.3 Assay efficiency

The PBS control was selected for determination of assay efficiency. After the rocking cycle of 24hrs, a 1mL aliquot of the PBS solution was dispensed into a clean 15mL Falcon tube and dried in vacuum for 24 hours. Subsequently, 10mL of ethanol was added to solubilize VD₃. Drying of the water was done in order to avoid spectral changes that would occur due to dilution of ethanol preventing proper interpolation of VDs concentration from the standard curve. After measuring VD₃ concentration, the efficiency was calculated by the following formula:

$$\text{Efficiency (\%)} = \frac{\text{VDs} + \text{VDf}}{\text{VDi}} \times 100\% \quad \text{Eq 3.2}$$

From the nine runs, two were statistical outliers and although have been shown in the runs plot, they were not included in the analysis. The runs plot shows that the assay is operating at $99.9 \pm 1.3\%$ efficiency.

This was not a good test of efficiency as PBS, which has low solubilization capacity, also has higher error by percentage. When this is incorporated in the equation above, the error is compounded and efficiencies greater

than 100% are observed. The expectation was that, due to the washing step (Step 6), there would be an underestimation of VDs which would result in consistently lower efficiencies than 100%. If the error was due to the initial dispensing of VDi aliquot in ethanol, then a deviation of 7%

from the nominal value of 1mg would be required to cause a supposed increase in

efficiency to 107%. This may not be instrumentally related but due to the clearing of ethanol from the Eppendorf tip. If the efficiency is truly 99.92% it would have to be owed to the lack of filtration and extraction steps that can result in loss of efficiency. A limitation of the efficiency test is that efficiency was only accounted for PBS and not in samples containing macromolecules. In order to give the assay rigour, one issue that must be explored in the future is the direct extraction of VD₃ from solutions containing macromolecules to determine efficiency.

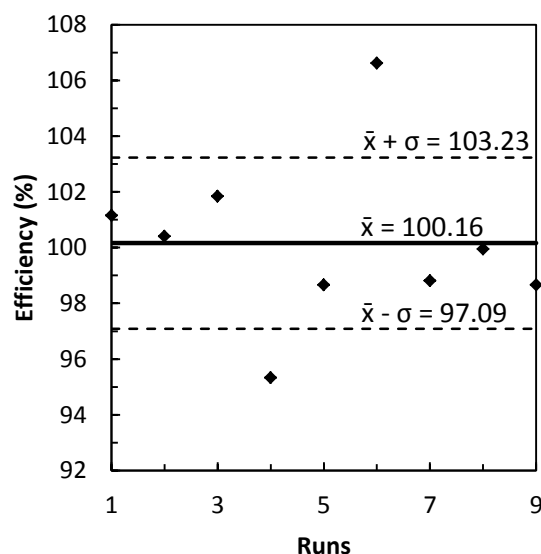


Figure 3.3 Efficiency of Indirect VD₃ Assay. Q-test for 4th and 6th run at 95% confidence showed that these values were not outliers.

3.5 Results and discussion

Due to the lack of data on lower and upper limits of VD_3 binding, five orders of magnitude of MM concentrations were tested - 100, 10, 1, 0.1, 0.01 g/L made in 0.1M, pH 6.7, PBS buffer solutions. Aliquots (10mL) were added to four VD_3 -primed tubes and after equilibration; VDf was measured to calculate VDs . Each concentration range of a specific MM was always conducted with three controls: EtOH, PBS buffer and RO water to verify VDi and to ensure control consistency to calculate amount of binding by the MM.

3.5.1 Kinetics

The main goal of the kinetics experiment was to establish the minimum time for attaining solubility equilibrium. Figure 3.4 shows a typical kinetics curve and although it is not a comprehensive plot of solubilization kinetics, it shows the important points for solubility of VD_3 in

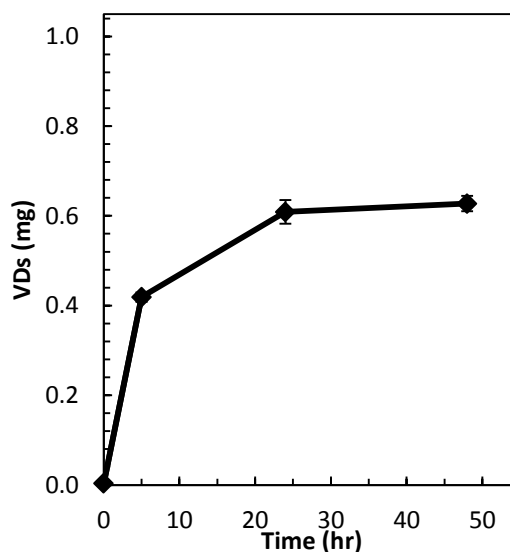


Figure 3.4 Kinetic solubility of VD_3 in 10g/L SCN-PBS solution; $n=4$.

a MM solution (SCN) over time. The first two points have a very high slope indicating a very rapid increase in solubility while the last two points have minimal slope indicating reduced rate of solubility. When the MM solution is added to a VD_3 -primed tube, some of the molecules from the adsorbed layer are solubilized in the PBS buffer. These solubilized molecules then bind onto the hydrophobic areas of the MM, thus shifting the equilibrium from the solvent to the MM. To maintain equilibrium, more solid dissolves in the solvent. VD_3 continues to solubilize until either its supply runs out or the binding

capacity of the macromolecule is reached. Qualitatively, in this case, the solubility equilibrium was established at ~24hrs, which is within 3% of the final value at 48hrs mark. Although there was a small increase, the distribution of VDs at 24hr and that of the 48hr mark makes either indistinguishable from the other. Thus, to minimize experimental time, and obtaining maximum solubility, 24hrs were deemed sufficient for all remaining solubility experiments.

3.5.2 Water and phosphate buffer

As expected, water (R/O) did not solubilize a significant amount of VD_3 0.014 ± 0.002 g/L (0.14 ± 0.02 mg in 10mL). Similarly, solubility of VD_3 in PBS was low and only slightly higher than that of pure water (R/O) at 0.018 ± 0.01 g/L (0.18 ± 0.01 mg in 10mL). Literature values vary on the solubility of VD_3 in water, however the range is between 0.0077 - 0.23 g/L¹⁰². Although it is tempting to state that the experiment was a success relative to the literature values and to the controls, the range itself was too wide and should be taken with considerable doubt. It is likely that compilation of VD_3 solubilities from different methods have yielded different results.

The online ALOGPS 2.1. ACLogS and AvLogS solubility algorithms calculated VD_3 solubility to be 1 μ M or 0.000385 g/L¹⁰³. The large difference between the theoretical and the experimental solubilities (~2 orders of magnitude) shows that a significant amount of work needs to be done in

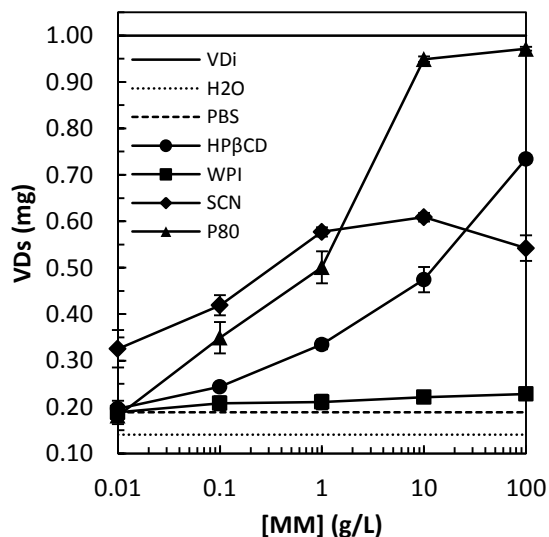


Figure 3.5 VD_3 -Macromolecule solubility; n=4. The solubility is interchangeable between mg and mg/ml by dividing by 10mL.

order to minimize the gap. A possible explanation for the difference was that small particulates may have broken off from the VD_3 layer in the Falcon tube to form a colloidal dispersion during mixing (Figure 3.6). Rather than measuring solubility alone, the solvation and dispersibility of VD_3 in PBS buffer were measured. The high VD_f value and very low amounts recovered for VDs in the PBS efficiency test indicated that

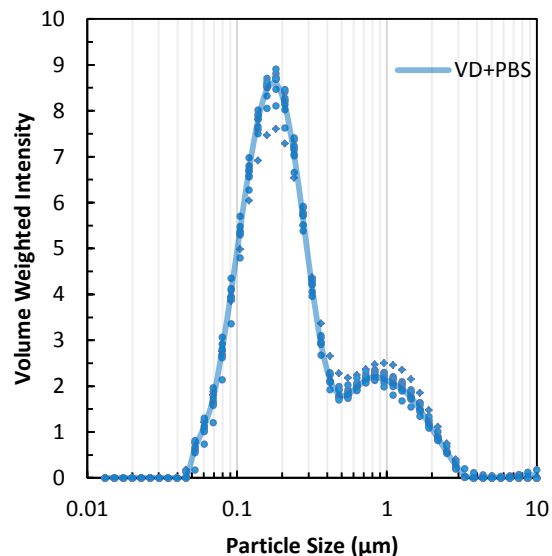


Figure 3.6 Particle size of PBS equilibration with VD_3 . $N=4$. Measurements of PBS-only yielded a flat line at 0 intensity at all particle sizes.

very low amounts of VD_3 were solubilized in PBS (Figure 3.5). As the particle size distribution of VD_3 in PBS ranged from 0.05-2.89 μm , this also strongly suggested dispersion rather than solubilization [as was observed in the P80 particle size distribution (Figure 3.10)]. For practical considerations, VD_3 solubility in the presence of the various macromolecules was compared to PBS and R/O results to determine binding. Binding was calculated based on the difference between VDs of PBS and VDs of macromolecule.

3.5.3 Whey protein isolate (native/denatured)

WPI has minimal binding to VD_3 at pH 6.7 (Figure 3.5). This work confirmed earlier findings where after cheese fortification, the collected whey contained no traces of VD_3 ⁶. There are two mechanisms that can account for this observation. As mentioned before, β -lactoglobulin is the one whey protein shown to have binding affinity towards VD_3 . The first of the two mechanisms is that, at pH 6.7, β -lactoglobulin is in dimeric form, which

prevents access to the hydrophobic ligand binding site on its surface. The second possible mechanism is based on the fact that, at pH 6.7, the residues of 85-90 involved in a connective loop between β -sheets are in a folded conformation closing off access to the external environment (see Figure 2.3).

If this study had been performed at pH 8.5, the so-called Tanford-transition of Glu9 from the calyx to the periphery of the protein would have permitted access of the non-polar molecules in solution to the central

calyx^{55, 56}. This would have registered as an increase in the amount of binding of VD_3 and macroscopically a decrease in the VD_3 layer at the bottom of the primed tube. If this hypothesis is correct, then during cheese processing, it is unlikely that any VD_3 binds to any of the whey proteins as the entire processing of the product is done at pH at or below 6.7. Thermal treatment of

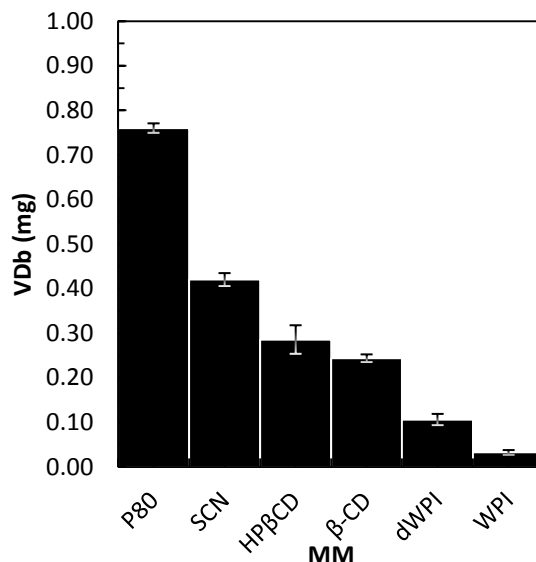


Figure 3.7 VD_3 -Macromolecule Binding at 10g/L MM and 1mg of VD_3 . Two additional molecules were tested for comparison β -CD and dWPI.

the WPI (dWPI) slightly increased binding observed, which was expected as internal hydrophobic moieties would be exposed⁵⁹. However, at pH 8, this has been shown to decrease the ability of β -lactoglobulin to bind VD_3 ^{11, 12}. At the native pH of milk 6.7, however, owing to Tanford transition, there was no access to the hydrophobic calyx of β -lactoglobulin and therefore an absence of VD_3 interaction *via* this mechanism. So denaturation of WPI and exposure of the hydrophobic groups has the relative effect of

slightly increasing VD₃ binding. Relative to native WPI, this increase was ~300% higher, though still very small compared to P80 and other MMs used in this study.

3.5.4 Cyclodextrin

Hydroxypropyl- β -Cyclodextrin (HP β CD) is an excellent solubilising agent for hydrophobic compounds and as it contains a hydrophobic core made of sugars. The binding curve for VD₃ in the presence of HP β CD showed a continual log scale increase. On the linear scale, there was an appreciable decrease in the change in binding especially above 1g/L of the macromolecule. This suggested that there may be another interaction involved. Literature data on HP β CD and VD₃ interaction indicates that a 2:1 binding may occur¹⁰⁴. That is to say at high concentrations, two molecules of HP β CD bind to one molecule of VD₃. Therefore, increasing the concentration of HP β CD did not lead to a linear increase in the amount of solubilized VD₃. Another cyclodextrin molecule, β -CD, solubilizes almost the same amount of VD₃ as HP β CD. This molecule is a cheaper alternative to HP β CD, and according Figure 3.7 also has the potential to be used for VD₃ fortification.

3.5.5 Sodium caseinate

Sodium caseinate has the potential to solubilise hydrophobic substances. The binding curve for SCN followed the expected trend up until 10g/L where a plateau was observed (0.61 ± 0.01 mg) for a concentration of 10g/L. At 100g/L SCN, a noticeable decrease was observed. To explain this trend, one has to recall that sodium caseinate in solution forms micelle-like structures maintained only by interactions between nonpolar residues on protein chains. The internal structure is a network of hydrophobic surfaces interacting with each other^{15, 24}. This network of hydrophobic surfaces occurs above the critical micelle concentration (CMC). Below the CMC, caseins exist as individual monomers

with high surface areas exposed to the water solution. As the monomeric SCN solution was introduced into the VD_3 primed tubes, due to the large solvent-accessible-surface-area per monomer (SASA/monomer), a large amount of VD_3 bound to the SCN. Increasing the concentration of SCN, still below CMC, increased the cumulative-hydrophobic-surface-area (CHSA) available and more

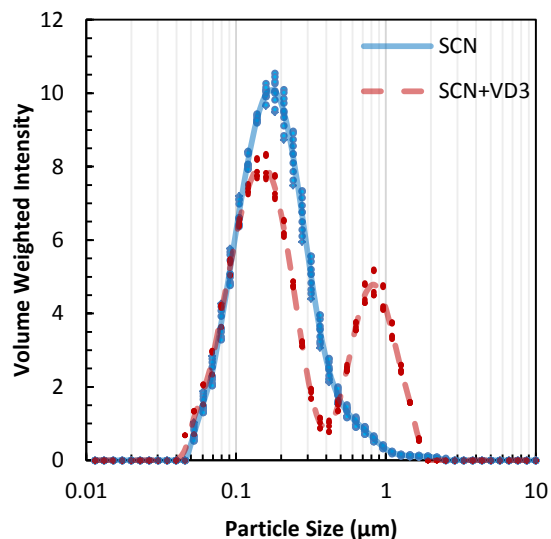


Figure 3.8 Sodium Caseinate micelles (10g/L SCN) before and after equilibration with VD_3 , N=4.

VD_3 molecules bound to the protein. Although there was still an increase in VD_3 binding, there was a decrease in the slope of binding. As the CMC was approached, there was some evidence of micelle formation. So, although the increase in concentration of SCN increased CHSA, the rise was relatively less due to a reduction in SASA/monomer as a result of micelle formation before CMC. Consequently, a decrease in the slope of binding was observed.

At the CMC ($\sim 1\text{g/L}$)¹⁰⁵, there was a small increase in binding corresponding to the increase in concentration of the macromolecule. However, there was also a continual decrease in the slope of binding, owing to the increase in preferential intramolecular hydrophobic interactions of casein monomers, respectively. The same was observed at 10g/L SCN. At 100g/L SCN, the hydrophobic interactions of caseins became so strong that although the concentration of casein increased, the CHSA was lower than at 10g/L, resulting in a decrease in VD_3 binding as well. Another explanation could be related to

experimental error associated with the SCN coating the VD₃ layer to prevent solubilization of VD₃ molecules. If that were true, it could be tested by adding the 100g/L solution of SCN to one VD₃-primed Falcon tube, and after 24 hrs, transferring the same solution to another VD₃-primed tube. If true, then some VD₃ would solubilize in the second tube as well as the first, as a significant portion of SCN molecules would have the potential to bind VD₃.

The highest VD₃ binding for SCN was at 10g/L of MM (0.042 ± 0.001 mg/mL). Past literature has reported VD₂ binding of 0.052 ± 0.003 mg/mL with 50g/L MM¹⁴. Accounting for molecular weight differences, the molar binding for the past literature was 0.131 ± 0.008 mM for VD₂ and for the current work of VD₃ binding, 0.109 ± 0.003 mM. There could be three reasons for the higher binding of VD₂ compared to VD₃. First, more SCN was used to bind VD₃. Second, ergocalciferol contains an additional methyl group which adds to its hydrophobic surface area and increases its propensity for binding to SCN. Third, the procedure used in the previous study utilized high pressure homogenization after addition of the vitamin to disrupt the casein-casein hydrophobic interactions effectively increasing the CHSA as well as SASA/monomer of the system.

Particle diameter of SCN micelles maintained a unimodal distribution ranging from 40nm to 1 μ m, with the average around 200 μ m (Figure 3.8) as have been reported multiple times in the literature^{14, 81, 106, 107}. An article by Semo et al., in 2005, reported that after vitamin binding, the reconstituted casein micelles did exhibit a small change in micelle diameter from 147-156nm. The possibility that casein micelles rearranged themselves to accommodate VD₃ is not difficult to conceptualize. The particle size of PBS solutions with solubilized VD₃ confirmed that there were microscopic VD₃ particles in the system

(Figure 3.6). So, the caseins in solution ‘decorated’ the surface of dispersed VD_3 leading to particles $\sim 1\mu m$ in size. Binding of some caseins to VD_3 reduced the number of unbound caseins resulting in a relative decrease in particle size intensity at $\sim 200nm$. The increase in average particle size due to VD_3 binding also increased the volume per particle and consequently, the decrease surface charge density. This can be observed with the measurement of ζ -potential which is directly proportional to surface charge density (Figure 3.9).

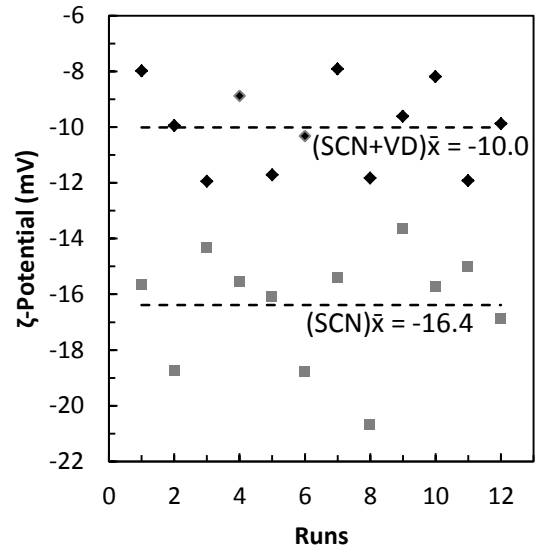


Figure 3.9 Runs plot of ζ -potential SCN before and after equilibration with VD. $N=4 \times 3$

In the study that looked at interaction between VD_2 and casein micelles, in addition to observing only a slight increase in particle size, there was no change in the morphology of the size distribution, which remained unimodal. The unimodal distribution was likely linked to procedural differences between the current and past studies. As high pressure homogenization was used after the dispersion step, the particle size distribution increased only slightly and did not exhibit bimodality. Although for the current study bimodality and a reduction in ζ -potential pointed to the reduced stability of the casein micelles, there is potential for further optimization of VD_3 and casein interaction through homogenization.

3.5.6 Polysorbate 80

As expected, P80 had the highest binding capacity for VD_3 . This was partially owing to its low CMC (0.012mM). At a very high concentration (100g/L), it was able to bind all of the available VD_3 adsorbed onto the tubes (1mg). Similar to SCN, there was a very small, but detectable increase in the particle size of P80

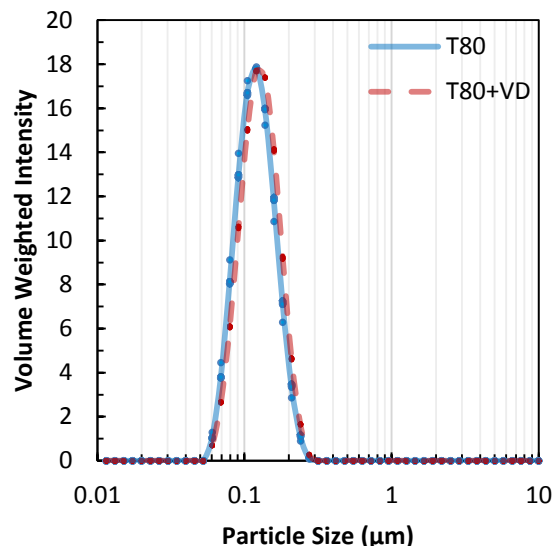


Figure 3.10 P80 micelles (10g/L) before and after equilibration with VD_3 . N=4.

micelles with VD_3 compared to the reference (Figure 3.10). The P80 micelles grew larger in response to accommodate VD_3 shifting the distribution from 120nm to 130nm.

3.6 Conclusions

The indirect solubility assay is a simple and efficient method to characterize binding of VD_3 , with macromolecules. It shows that although P80 is excellent at binding to VD_3 , other fortifying agents such as SCN and $\text{HP}\beta\text{CD}$ may also be used for VD_3 fortification and if optimized may protect it from the processing environment encountered. It also shows that WPI is unable to bind to VD_3 at all, its denaturation leads to a small increase in binding. The molecular reason for this observation is related to the Tanford transition of 85-90 amino acid residues, which prevents access to the hydrophobic barrel of the β -lactoglobulin whey protein. Finally casein micelles in SCN solutions can bind to VD_3 . Future work could be on optimization and localization of bound VD_3 in the casein micelle.

3.7 Acknowledgements

The authors of this work would like to thank Drs. Frances Sharom, Dr. Rickey Yada, Brian Bryska, and Vladimir Bamm for their help in acquiring access to and training on the ITC. Sofija Katic is acknowledged for her help with fluorescence expertise. Moumita Ray and Renuka Gupta are acknowledged for their help with Malvern Mastersizer. Miriam de Jong is acknowledged for her help with access to the VWR rocker platform. The sponsors of this work were NSERC, AFMNet and Dairy Farmers of Canada.

Chapter 4: Structure of β -casein f(1-25)

4.1 Abstract

The secondary structure of β -casein phosphopeptide 1-25 (β -CPP) is currently under debate due to its intrinsic disorder. A series of replica simulations (1, 10, 100 replicas) with cumulative time of 1.0 μ s was performed in an explicit solvent system (1atm, 298K) with 0.1M ionic strength. Prior to analysis quality assurance was performed to ensure equilibrium and absence of non-physical effects. Efficiency of sampling the conformational space was carried out by using H-bond analysis, which showed highest sampling with 100 replicas (94.6% of the theoretical maximum 527). Structural features, especially interactions between negatively charged residues E₂ - S₁₅ and S₁₇ - R₂₅ were responsible for incomplete sampling of conformational space owing to electrostatic repulsive forces. From the simulation, trajectory radius of gyration, H-bonding, Ramachandran plot, and DSSP secondary structure were probabilistically analysed. It was observed that β -CPP adopts a highly compact structure (Rg 0.75-1.1nm, peak: 0.84nm), the conformation of which was governed mostly by local interactions rather than long range interactions. With the exception of terminal residues and Glycine (G₁₀), most amino acids preferred bend-type structures, with a lower probability for turns due to steric repulsion. Low probability of helix and β -sheet structures was also observed. Glycine was found primarily in loop-type structures due to its conformational flexibility. Furthermore, there was significant propensity towards angles observed in PPII, extended β -sheet, β -turn and α -helical structures. Results of the simulation were in good agreement with simulations of other disordered peptides and well as experimental structural data of β -CPP. Future studies with multiple replica sampling will be directed towards structural simulations of β -CPP in various physicochemical conditions, membrane simulations to understand structure in presence of various divalent cations, as well as testing the method with other parts of β -casein protein to observe binding to VD₃.

4.2 Introduction

β -casein is an abundant protein in milk composed of 209 amino acids. It is an amphipathic protein given its hydrophobic C-terminus¹⁰⁸, and the hydrophilic N-terminus constituting of highly-charged phosphoserine residues in the fragment (1-25)⁷¹. The N-terminus contains four phosphoserines that determine the net charge of the protein. These phosphoserines are responsible for the formation of calcium phosphate nanoclusters through electrostatic sequestration¹⁰⁹. The formation of these nanoclusters prevents ectopic calcification in mammary tissues that have high calcium flux¹⁹.

In 1981, the School of Dental Science at University of Melbourne in Australia, showed that milk-derived foods can help prevent dental caries. Further investigation showed that specific casein peptides were responsible for this anticariogenic property of milk^{28, 83, 84}. These peptides contained special phosphorylated serine residues which acted as nucleation sites for calcium and phosphate binding. Since this discovery, this technology has been patented and licensed under the name Recaldent¹¹⁰.

One of those anticariogenic peptides is the N-terminus fragment of β -casein fragment 1-25. This fragment (β -CPP) is nutritionally active and has been effective in treating and preventing oral diseases⁸³. It is able to mitigate biofilm formation and concomitantly favours nucleation of possibly apatitic calcium phosphates¹¹¹. It has been also implicated in improving intestinal absorption of calcium through *in vitro* studies¹¹². The high degree of calcium absorption in caseins is governed by the sequence flanking the multi-site phosphorylation Ser(P)₃-Glu₂ motif^{84,112}. For example, if the Arg1-Glu4 segment is removed or inverted with Val8-Glu11, or if the acidic motif is removed, any of these combinations prevents calcium absorption¹¹². Speculation is that the mechanism of

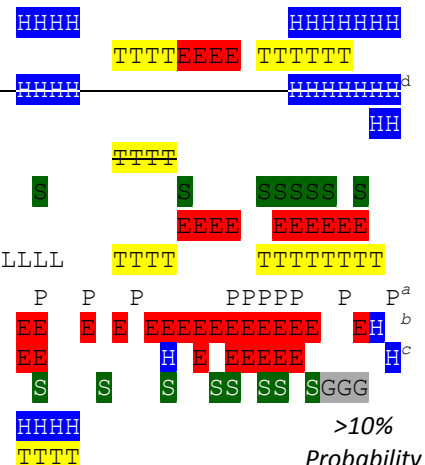
binding to the plasma membrane of intestinal epithelial cells is mediated by electrostatic interaction of the flanking chains of the phosphoserine motif¹¹². The binding properties of β -CPP, its ability to form soluble calcium phosphate complexes²⁷ and its resilience against proteolysis¹¹³ supports its function as a calcium carrier.

About 2.5% of infants under the age of three cannot use the calcium-carrying function of the β -CPP peptide as it elicits an allergic response⁸⁵. The N-terminus region of the calcium carrier peptide is a major epitope for IgE/IgG binding f(1-16) and is known as *Bos d 8* in allergen databases^{85, 114, 115}. Subsequent to β -CPP binding to the Fab region of the IgE, a cascade of biochemical processes can cause hives in mild cases and anaphylaxis in severe cases⁸⁶. Of the major antigenic regions of β -casein, the N-terminus of β -CPP (1-16) is the most frequently-recognized epitope⁸⁵. Whether the mode of epitope recognition is shape or simply charge complementary, is unknown. The high net charge of $-13e$, the prevalence of proline, and of glutamic acid residues add to the evidence that β -CPP is structurally disordered.

Various studies have attempted to elucidate β -CPP conformation, using circular dichroism (CD)^{116, 117}, NMR^{28, 118, 119}, FTIR¹¹⁶, Raman techniques^{22, 120-122}, sequence-based prediction methods and molecular modeling techniques^{116, 123, 124}. Table 4.1 is a compilation of proposed secondary structures over the last 30 years for this fragment. Few regions are consistently predicted, but overall there is severe disagreement amongst the various structural studies. Such disagreements in experimental structure determination techniques (CF, CD, Raman, FTIR) arise in large part from the associated difficulties in data processing databases, which are usually based on pre-existing globular protein structure characterizations. Such characterizations are not applicable to unfolded proteins

or peptides such as β -casein or β -CPP as they bias proposed conformations towards folded, globular, structures and away from relatively unstable conformations like the polyproline-II-helix (PPII) and extended β -strands²⁴. Using NMR for structure determination of IDPs is also complicated due to their dynamics which result in multiple overlapping peaks, deconvolution of which is not always possible²¹.

Table 4.1: DSSP representation of β -CPP secondary structure. *Experimental methods* CD: Circular Dichroism; NMR: Nuclear Magnetic Resonance. *Sequence-based Prediction methods* CF: Chou-Fasman¹²⁵; GR: Garnier, Osguthorpe & Robson¹²⁶, RP: Robson & Pain¹²⁷; SE: Schiffer & Edmundson¹²⁸; VIL: Lim VI¹²⁹; WK: Wu & Kabat¹³⁰. *Theoretical methods* EM: Energy Minimization; MD: Molecular dynamics.

(G: 3 ₁₀ , H: α , I: π) helix, B: β -bridge, E: Extended OR β -sheet, T: Turn, S: Bend, L: Loop, P: PPII Helix				
1°	RELEELNVPGEIVE ¹⁰ SL ²⁰ SS ²⁰ SEESITR	Technique	Reference	Year
2°		CD+CF,GR,RP,SE,WK	* ^{131, 132}	1975/81
		CD+CF,VIL	* ¹³³	1984
		CD ^d	¹¹⁷	1988
		CF	* ²³	1987
		NMR ^e	¹¹⁸	1991
		<i>In vacuo</i> EM	* ²⁹	1993
		NMR	¹¹⁹	1993
		NMR	²⁸	2001
		CF		
		Chemical Shift	¹¹⁶	2002
		NMR Shifts		
		MD		
		<i>In vacuo</i> MD	¹²⁴	2006

Strikethrough indicates absence of specified structure. Asterisk (*) indicates structure based on entire β -casein protein. Helix formation observed in >40% 2,2,2-trifluoroethanol^d. Type II β turn propensity, but not all connectivities found^e.

Molecular dynamics (MD) simulation has been hailed as the ‘holy grail’ of protein structure determination²⁵. Intrinsically disordered proteins (IDPs) are difficult to study structurally due to their conformational heterogeneity²¹, which results in multiple degenerate conformational states – a characteristic specific for disordered proteins²¹. As computer simulations are not limited by these experimental constraints, they can provide

atomic resolution of IDPs²⁶. Since the first energy minimized structure of β cn was published²⁹, numerous simulations have appeared on this protein. Wang *et al.* (1994) studied the conformation of a peptide of β cn f(93-103) to verify the predictions made by Holt & Sawyer in 1993 regarding the presence of a α -helix in that region¹⁸. Later, Yasar *et al.* (2006) investigated the conformation of three short peptides of β cn predicted to have α -helical structure by Kumosinski *et al.*'s energy minimized model²⁹. By performing multicanonical *in vacuo* simulation, Yasar *et al.* showed that the three peptides f(1-6, 35-40, 43-48) exhibited a high probability (>10%) towards forming α -helices and turns as shown by Kumosinski's model¹²⁴. Farrell *et al.* (2002) also performed a simulation on β cn(1-25) to determine the relative mobility of the phosphorylated *versus* dephosphorylated peptide in an 200ps explicit solvent MD simulation¹¹⁶. Other simulations by Cassiano *et al.* (2003) determined the critical role of phosphoserine residues on interfacial properties of β cn peptides¹³⁴.

The aim of this work was to determine the structural ensemble of the β -CPP peptide through MD simulation to i) verify/eliminate the various secondary structures observed in the literature; ii) understand structure-function correlations and, iii) establish a method for studying small disordered peptides through MD simulation.

4.3 Materials and methods:

4.3.1 Molecular dynamics Set-up

The primary sequence of β -CPP was obtained from DisProt (RELEELNVPGEIVESLSSEESITR) from which a PDB file was generated through Pymol (DeLano Scientific; <http://www.pymol.org>). The structure was modified to add a phosphate groups with a charge of $-2e$ at the known loci of phosphorylation (Amino acids: 15, 17, 18, and 19 from the N-terminus)^{16,68}. The PDB file was then imported into GROMACS 4.5.1¹³⁵ using a modified GROMOS 43a force field for phosphoserine groups. This is a united-atom force field that combines non-polar hydrogen atoms onto the bonded adjacent atoms and was parameterized for protein simulations for all 20 common amino acids as well as their phosphorylated derivatives¹³⁶. Box assignment of dodecahedron at an edge to peptide distance of 1.4nm was used as the simulation box to solvate the protein with 3249 simple point charge (spc) water molecules. The high $-13e$ charge was neutralized with the addition of 20 Na^+ and 7 Cl^- ions. Excess sodium and chloride ions ensured $\sim 0.1\text{M}$ ionic strength. After energy minimization through steepest descent and conjugate gradient in tandem, in order to prevent bad van der Waal contacts, a 2ps isothermal-isobaric (298K, 1atm) ensemble simulation (V-Rescale temperature coupling and Berendsen isotropic molecule-based-scaling for pressure¹³⁷) was performed to ensure constant temperature and pressure. Position restraint simulation was also performed for 10ps to allow water to equilibrate to the peptide. For unrestrained simulation the electrostatic interactions were evaluated by Particle-Mesh Ewald method¹³⁵ with a charge grid spacing of approximately 1.4nm while the Lennard-Jones interactions were evaluated using a 1.4nm cutoff.¹³⁵

4.3.2 Quality assurance

To determine whether the starting structure made a difference on the equilibrium conformations, three different simulation versions each totalling 1.0 μ s with 0.1ps sampling, were performed. These included one single simulation of 1.0 μ s, 10 simulations of 0.1 μ s, and 100 simulations of 0.01 μ s. The initial starting structure of each simulation was selected from a pool of one thousand different conformations obtained from fifty 10ns NVT simulations conducted at 1000K with sampling every 0.5ns. Equilibrium for each set of simulations (1, 10, or the 100 series) was ensured by zero drift in radius of gyration (R_g) which was assumed to be the structural parameter with the longest equilibration time. The initial portion containing R_g decay was removed to avoid convolution of kinetic changes with equilibrium properties of the system. The simulation trajectories were extended to compensate for the removed kinetic segment. In addition, the lack of interactions between adjacent periodic images (minimum distance always >1.4nm) ensured no contact throughout the simulation, which could have led to nonphysical effects. Finally, the total number of hydrogen bonds (H-bonds; distance between N-H and C=O \leq 0.35nm) was selected as an equilibrium-convergence-metric for conformation.

4.3.3 Simulation analysis

The unrestrained simulation was analyzed using the in-built tools provided with GROMACS for R_g , H-bonds, Ramachandran plot, and secondary structure analysis (using DSSP)¹³⁸. All analyses were based on probability.

4.4 Results and discussion

4.4.1 Ensemble equilibration

Relaxation time was observed by plotting radii of gyration from the 100 series over simulation time (Figure 4.1).

The radii were initially spread over a range of 0.8-1.6nm, but as they were simulated the range shifted and became smaller 0.75-1.1nm. The initially large spread was due to the high temperature simulation that generated the hundred different structures. Although in reality at temperatures of 1000K, the peptide would break apart, using constrained simulation where bonds do not break and only conformation space is sampled is a common simulation technique to generate different starting structures²⁶.

At high temperatures (1000K in this simulation), the system has enough energy to overcome most conformational barriers present at 298K. As a result, the rugged conformation energy landscape becomes smooth and the conformationally-restrained peptide can diffuse homogeneously in any direction along the radius of gyration energy landscape. As the temperature is lowered back to 298K, the ruggedness of the landscape

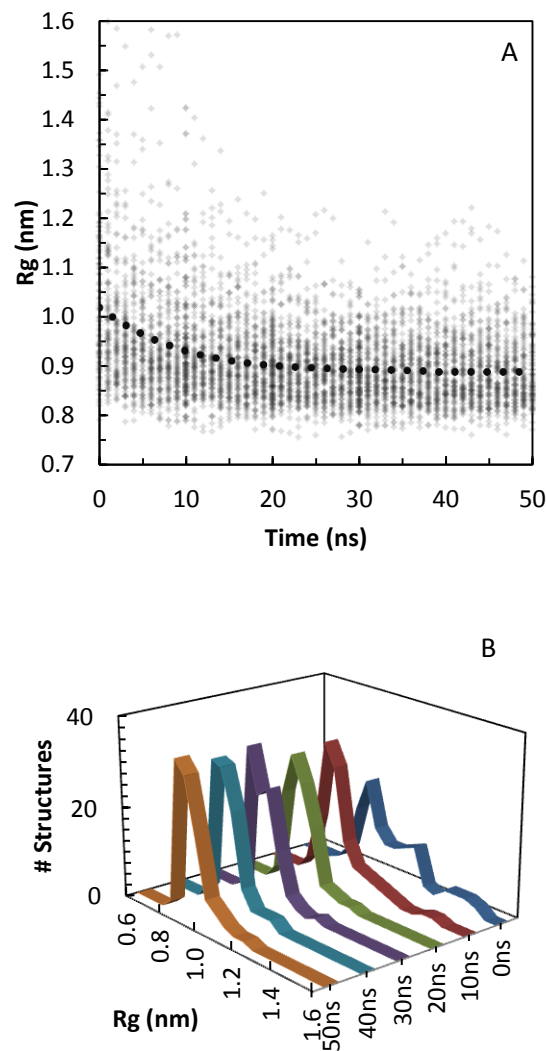


Figure 4.1 Radius of gyration for 100 replica series over time (A). The dotted trendline is a 4th order polynomial. (B) R_g of the various replicas over time. to determine adequate equilibration time.

reappears and each replica falls into a location where it can either stay or ‘slide’ down to attain a local energy minimum. As the peptide replicas slide, they adopt an energetically favourable and more compact conformation. Overall, this has the effect of reducing the spread of radii of gyration to 0.75-1.1nm. Monitoring R_g over the simulation trajectory also allows estimation of the ensemble equilibrium relaxation time.

Figure 4.1 shows that in the 50ns of simulation time of 100 replica series, after 30ns, the distribution for the radii of gyration, although dynamic, had a constant average and minimal drift. In addition, Figure 4.1b shows the R_g of the 100 replicas over simulation time. At the onset of the 100 replica simulation, the distribution morphology was highly variable. As the simulation proceeded, the ensemble of β -CPP moved towards an overall energy minimum and a lognormal-shaped distribution appeared. This occurred after 30ns of simulation time. For all replica series, the initial 30ns were removed and compensated at the end of the simulation. A cumulative simulation time of 1.0 μ s under equilibrium conditions were used for trajectory analysed.

4.4.2 Simulation performance

After determination of equilibration time, the next prerequisite of MD simulations prior to analysis was to determine the quality of simulation. This was accomplished by a convergence metric which shows the performance of the simulation. For known disordered peptides like β -CPP, where the ultimate goal is to understand the conformation ensemble instead of acquiring convergence to a singular structure, knowing the quality of simulation becomes very important. For the purposes of this work, intramolecular H-bonds were used to as the convergence metric as has been previously shown²⁶. For any peptide of N amino acids and P prolines, the maximum number of

hydrogen bonds that can be observed in statistical ensemble can be obtained by the formula:

$$Total\ bonds = N^2 - N - 2(N - 1) - N \times P \quad \text{Eq 4.1}$$

The first N^2 term calculates the total permutations, the second term subtracts the diagonal *i.e.*, H-bonds with self, and the third term removes the sterically-improbable H-bonds between adjacent amino acids. Finally $-N \times P$ removes the N-H instances that do not exist on prolines (P). For a graphical picture of H-bond contact map and to understand where this formula comes from, see Figure 4.5. Given that the β -CPP contains 25 residues, one of which is a proline, 527 H-bonds can theoretically be sampled by a statistical ensemble of β -CPP peptides. For consistency, H-bonds will be represented from acceptor to donor residues ($C=O \rightarrow N-H$).

Figure 4.2 shows the cumulative H-bonds observed during the course of the simulations.

The number of H-bonds increases with increasing replicas from 146 to 368 and the highest occurring at 499 H-bonds for the 100 replica series. About 94.7% of the theoretical numbers of H-bonds were observed with the 100 replica series owing to increased sampling. It is also interesting that even one long simulation is able to sample 27.7% of the total H-bonds. This was highly unexpected, as

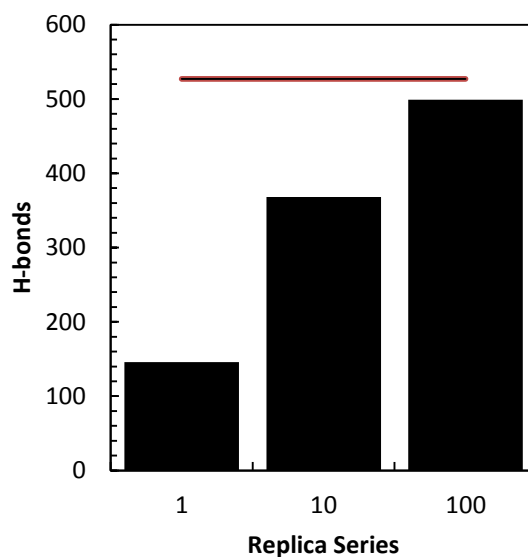


Figure 4.2 Number of H-bonds observed during the course of simulation. The line above the graph indicates the theoretical maximum of H-bonds possible with β -CPP.

from the H-bond contact maps (Figure 4.5) the probability of observing many H-bond contacts is low in a singular trajectory. The large number of H-bonds may simply be due to local instead of non-local interactions. Every H-bond interaction takes place between two residues, a donor and an acceptor residue. Each residue has two functional groups and two immediate neighbours. So a non-terminal residue can have eight immediate local interactions, most of which may simply be transient. To sample local H-bonds, the peptide would have to have some degree of mobility, which for β -CPP may be provided by intramolecular electrostatic repulsions. The transiency of most local H-bonds is apparent from the sparsely populated contact probability map, but in H-bond ($S_{24} \rightarrow S_{17}$) local sampling of H-bond can be observed (Figure 4.5A).

4.4.3 Radius of gyration

Probabilistic analysis of Rg yielded a bimodal distribution for 1 and 10 replica series and an overall generalized lognormal distribution for 100 replica series. The peak of the 100 series was localized at 0.84nm. For the 1 replica series, two different conformations were observed. The absence of any significant radii of gyration in between the two modalities indicates that the one replica structure is vacillating between two different conformations. These two conformations have been shown in Figure 4.3 as structures A and B. Both structures are primarily stabilized by electrostatic interactions between the positive, terminal Arginine residues the highly negative phosphoserine residues.

The main difference between the two structures is that the smaller structure arises from hydrogen-bonding between $E_4 \rightarrow S_{24}$ and $V_8 \rightarrow S_{24}$, which pulls the residues between E_4 and V_8 towards the positively-charged C-terminus thus reducing Rg. Though this smaller Rg suggests a more stable peptide structure, this is not likely due to the inherently high charge. This is one reason why in the 100 replica series, the structures with low Rg values had a low probability region from 0.75-0.80nm. Having a small Rg for β -CPP would require precise hydrogen and electrostatic bonding

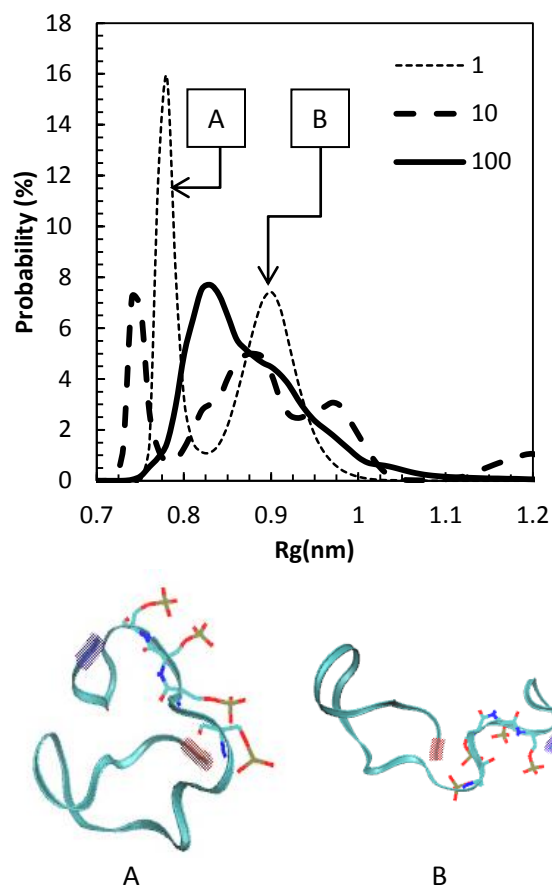


Figure 4.3 Rg of 1, 10 and 100 replica series of β -CPP peptide as well as corresponding structures. The initial red portion on structures identifies the N-terminus, while the blue end is the C-terminus. Phosphoserine structures are shown as stick models.

structure. Since the precise electrostatic bonding structure can be formed by only a few stable structures, it is thus not a favourable β -CPP conformation relative to the ensemble, due to tight packing of negatively charged residues. However, for the 1 replica simulation, the tightly packed structure Figure 4.3A was significantly more stable than the larger structure Figure 4.3B as shown by the probability distribution. The 100 replica Rg probability peaks at 0.84nm before tapering at higher Rg (Figure 4.3). The smooth reduction in Rg population at the other extreme could be rationalized the opposite way. If a peptide were to have a large Rg, due to the charges as well as potential to form

hydrogen bonds present the system would favour formation of smaller sized structures. As a result, after the initial compact structural conformation, the increase in size is a continuum which corresponds to a rapidly decreasing population.

The sum of all replicas tended towards asymmetric normal distributions, as already observed for IDPs^{21, 26}. These results strongly suggested that for β -CPP peptide, there are multiple, degenerate energy minima in which the peptide can exist. The most probable R_g of 0.87nm was compared with bioinformatics data analysis of 1000 globular proteins in the PDB database, using the following relationship between the number of amino acids (N) and their corresponding R_g :¹³⁹:

$$R_g = 0.395 \times N^{\frac{3}{5}} + 7.257 \quad \text{Eq 4.2}$$

This equation, when used for β -CPP peptide predicts an R_g value of 1.0 ± 0.07 nm, assuming the peptide is in a globular state, which, unfortunately is an overestimation even considering the error of regression. An enhanced sampling simulation of an IDP of 35 amino acids peptide (GVPGV)₇ in the literature exhibited a smaller R_g located at ~ 0.79 nm. Although the peptide was larger, the lack of negative residues presumably allowed closer packing and a reduced R_g compared to β -CPP. A very detailed SAXS and SANS study by Holt *et al.* 1998 measured the R_g of the β -CPP peptide when complexed with calcium phosphate. Through model calculations, it was discovered that the R_g of β -CPP in the complexed state was 1.85 ± 0.05 nm¹⁴⁰. However, the R_g plot from the simulation indicated a very low probability for any value above 1.1nm.

If this were found to be the case in a similar set of simulations of β -CPP with Ca^{2+} ions present, it would indicate an open conformation that occurs upon complexation of the

peptide with calcium phosphate. The open conformation may act as a brush that increases the local entropy to prevent further growth of the nanoclusters similar to the role of κ -casein in maintaining steric stability to casein micelles^{17, 66, 71, 141}. Given a sufficient concentration of the peptide relative to CaHPO_4 , it would also prevent mineral conjugation with other nanoclusters through steric repulsion. The mechanism governing the conformation change could simply be charge neutralization of phosphoseryl groups by the calcium phosphate salt which would displace the amino acids formally interacting with SerP, to allow their association with water and other ions in solution.

4.4.4 Hydrogen Bonding

Hydrogen bonds in proteins are very important as they are involved with structural integrity of proteins¹⁴³. The protein backbone is made from a series of peptide bonds between adjacent residues that determine the primary sequence of a protein. Peptide bonds form by a dehydration reaction between a carboxylic acid on the first, and a primary amine on the second, amino acid (Figure 4.4A)¹⁴². The peptide bond that forms has a partial double bond character as the electronegative oxygen can pull electron density from the

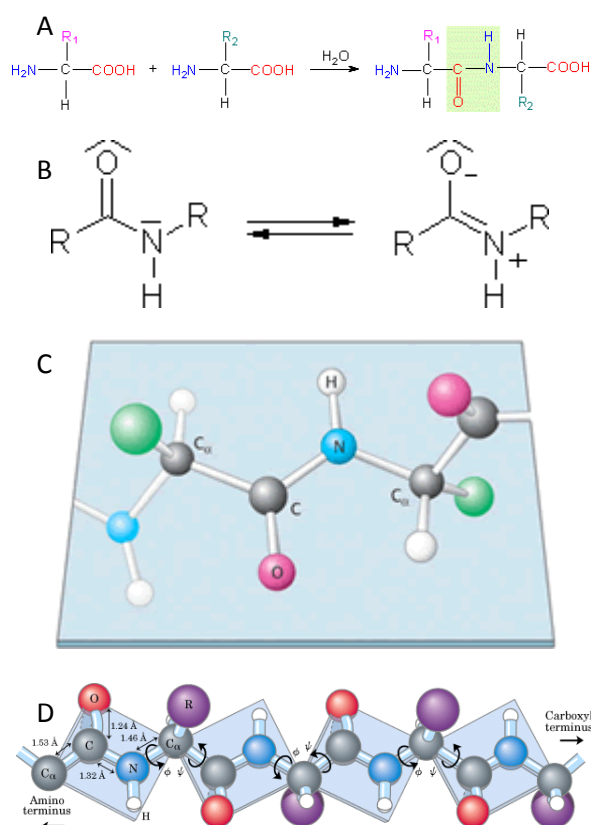


Figure 4.4 Formation and characteristics of a peptide bond. (A) Formation of peptide bond, (B) double bond character, (C) planarity, (D) allowable rotations of C^α -N and C^α -C(O) bond.^{142, 143}

nitrogen through its connected carbon (Figure 4.4B)¹⁴³. As a result, the amido group atoms H-N-C-O of the peptide bond all lie in the same plane (Figure 4.4C). The double bond character of the peptide bond has a marked effect on rigidity of the polypeptide chain and consequently on the folding of the polypeptide chain as well¹⁴². Rotations are only allowed between C_α-N (φ angle) and C_α-C(O) (ψ) angle (Figure 4.4D). Plots of C_α-N (φ angle) vs. C_α-C(O) (ψ) angles are known as Ramachandran plots and can give insights into the propensity of angles toward a particular secondary structure. The peptide backbone can participate in hydrogen bond formation, which occurs as a result of donation of electrons from the carboxyl group (C=O) to the amino group (N-H)^{21, 142}.

Figure 4.5 shows acceptor atoms (C=O) on the x-axis and donor atoms (N-H) on the y-axis of the same individual residue. The residues are shown from the N→C terminus. The plot was discretized into indices corresponding to each donor and acceptor residues. Along the trajectory, the number of times each index was encountered, it was tabulated. At the end of the simulation, the number of H-bonds in each index was normalized to the total number of H-bonds to obtain the probability map. Recall that for H-bonds plot the center diagonal, the adjacent residues, and the N-H groups of prolines are removed. There is asymmetry in the plot along the diagonal because the two axes correspond to two different groups of atoms on each residue (N-H, C=O). The top left side of the map from the diagonal is dedicated for N-H group interactions with C=O on the peptide backbone whereas the lower right side is dedicated to C=O and their interactions with N-H groups. Furthermore, most hydrogen bonding interactions have directionality. Mathematically this translates into the donor residue number almost always being larger than the acceptor

residue. When the donor residue number is smaller, it leads to rare secondary structures that are known as right-handed (*e.g.*, right-handed helix).¹⁴⁴

The first notable feature in all H-bond contact maps is that the maximum probability of each map is different as noted on the colour bar. With increasing number of replicas, the maximum probability decreases from 0.0587 to 0.0252. As a result, multiple replicas approach the statistical ensemble and sample a greater range of the conformation space. Due to multiple starting conformations, the population of H-bonds sampled greatly increases and ‘lights up’ different areas of the contact map. As a result, the H-bond contact probability spreads over several contacts instead of being restricted to one set of contacts due to an energy barrier. The most favourable contact has an overall reduced probability which is more representative of the ensemble. This does not limit the analysis as the contact map with reduced peak probability can be scaled to the maximum probability to show areas that have been sampled during the simulation trajectory.

One important observation common to all contact maps is the presence of the third diagonal from the centre towards the donor N-H groups. This diagonal is heavily populated with the highest probability contacts between i (C=O) and $i+3$ (N-H) residues. According to Table 4.2, this type of H-bonding refers to 3_{10} -helix. These helices are rare due to their tight angular requirements. It is likely that the simple H-bonding definition (0.35nm cutoff between donor N-H and residue C=O) overestimates the number of H-bonds. In order to obtain the best measure of H-bonds the definition would allow for 30° along the N-H—O bonds, a 0.25nm H to O distance, and 0.35nm N to C distance cutoff¹³⁵. Therefore to improve upon this work, the ideal definition has to be taken into account.

Even with overestimation of H-bonds, there are some reasonable conclusions that can be deduced based on the 100 replica simulation shown in Figure 4.5C. There are three regions of very high contact probability along the backbone ($E_5 \rightarrow N_7$, $L_6 \rightarrow V_8$; $E_{11} \rightarrow V_{13}$; $E_{21} \rightarrow I_{23}$, $I_{23} \rightarrow R_{25}$). Usually this type of a pattern denotes the presence of β -turns¹⁴⁵⁻¹⁴⁷. Another implication of this result is that local interactions along the backbone of the β -CPP are highly favoured rather than non-local interactions most likely due to the presence of charges along the backbone of the peptide. Due to increased sampling, two important features appear in the contact map of the 100 replica series. First, the N-H side of the map is more populated than C=O side, signifying a higher degree of H-bond interactions for N-H than their neighbouring carboxyl (C=O) counterpart. As mentioned before this is due to directionality of H-bonds¹⁴⁴. Second, the regions of high negative charge from residue 1-14 and residues 15-19 have a low probability of interaction because of electrostatic repulsion.

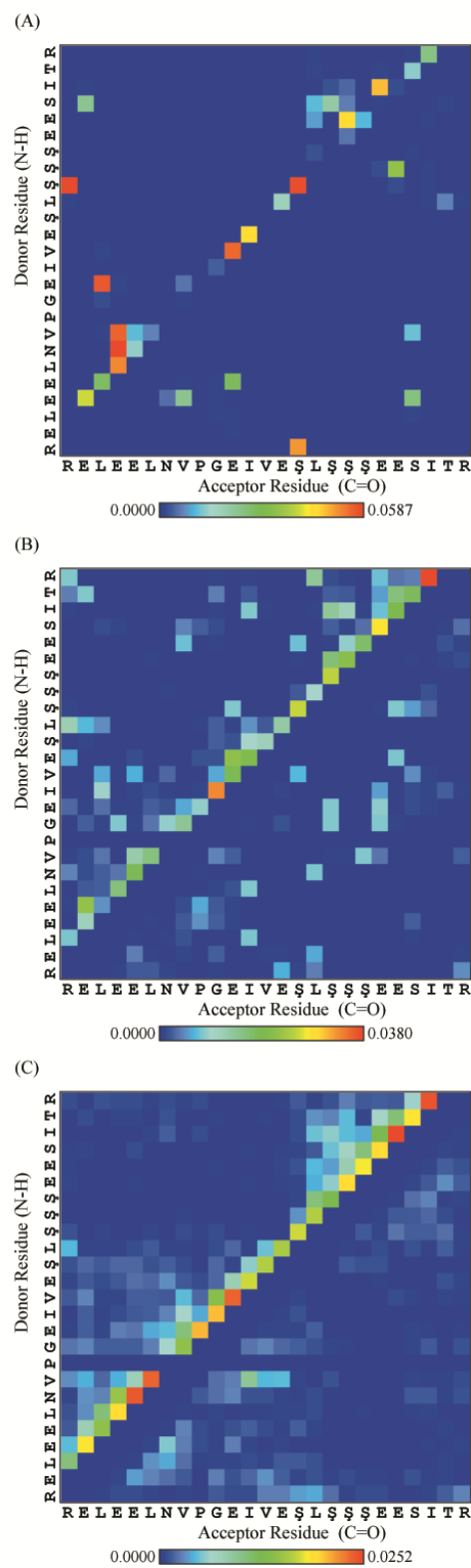


Figure 4.5 H-bond contact map for 1 (A), 10 (B) and 100 (C) replica series. Color indicates probability.

4.4.5 Secondary Structure

Some studies have emphasized the presence of secondary structure in β -CPP peptide localized in particular residues¹¹⁶ while others have hypothesized that the peptide maintains an overall worm-like conformation defined as rheomorphic where local structure may exist, but it is not localized to a series of residues²⁴. This is where the DSSP probability plot (Figure 4.6) becomes very useful as it can isolate secondary structures during the simulation along with visualization of their drift. The DSSP algorithm takes into account hydrogen bonds as a primary and rotation angle as a secondary measure for structure assignment. The elementary units of higher order structures are ‘turns’ and ‘bridges’. Repeating turns develop into helices while repeating bridges become ‘ladders’ (commonly known as extended structure), and repeating conjugated ladders form ‘sheets’. Regions of high curvature without hydrogen bonding are classified as ‘bends’. Lastly, a catchall term for every other geometry is called a ‘loop’^{138, 138}. Table 4.2 identifies the eight different secondary structures identified by the DSSP software along with their mechanism of formation, and simplified molecular structure.

Similar to the H-bond analysis, increasing the number of replicas from 1 to a 100 increases the quality of the results (Figure 4.5). Each structure observed in DSSP plot does not occur by virtue of one residue. For the purposes of generating the plot, the secondary structure each amino acid residue experiences during the simulation, was isolated and the probability was calculated based on the number of times it was observed in the simulation trajectory.

Table 4.2 DSSP characterization of secondary structures, mechanism and molecular structure. X identifies the residue for which the calculation is taking place. The punctuations “!” and “.” mark hydrogen bonds. In cases where X is not marked, a window of select size (defined by)is used to identify the presence of a structure . ¹⁴⁴ The 3-, 4- and 5- turns are characterized as a turn in the final DSSP plot, but if stacking of turns leads to different helical structures, this is identified separately.			
	SS	Mechanism	Molecular Structure
G	3 ₁₀ -Helix	H-bond	2x (3-Turn)
H	α-Helix	H-bond	2x (4-Turn)
I	π-Helix	H-bond	2x (5-Turn)
B	Bridge • Parallel	H-bond	<pre> X -N-C-C-N-C-C-N-C-C- H O H O H O \ . . / \ . / \ / H O H O H O -N-C-C-N-C-C-N-C-C- </pre>
	• Antiparallel	H-bond	<pre> X -N-C-C-N-C-C-N-C-C- H O H O H O . ! ! . . ! ! . . ! ! . O H O H O H -C-C-N-C-C-N-C-C-N- </pre>
E	β-Sheet Extended	H-bond	2x (Bridge)
T	Turn • 3-Turn	H-bond	<pre> -N-C-C-N-C-C-N-C-C-N-C-C- H O H O H O H O >-----< </pre>
	• 4-Turn	H-bond	<pre> -N-C-C-N-C-C-N-C-C-N-C-C-N-C-C- H O H O H O H O H O >-----< </pre>
	• 5-Turn	H-bond	<pre> -N-C-C-N-C-C-N-C-C-N-C-C-N-C-C-N-C-C- H O H O H O H O H O H O >-----< </pre>
S	Bend	Direction change >70°	<pre> X i-2 i-1 i i+1 i+2 -N-C-C-N-C-C-N-C-C-N-C-C-N-C-C- !----->!----->! </pre>
L	Loop		Not G, H, I, B, E, T, S

As an example, glycine (G₁₀) was found in a loop structure 70% of the time, 25% of the time in a bend, ~4% in an extended structure and 1% in a turn or a β-bridge. So the colour coding of the plot corresponds to that probability which shows yellow colour

(~0.75) at the loop, light blue (~0.25) at a bend, blue (~4%) at extended or the β -sheet structure, and dark blue (~1%) at β -bridge and turns. The colour and intensity, which is based on the rainbow lookup table, indicates the probability of SS occurrence for an individual residue. The abbreviations on the y-axis are for DSSP structure characterisation and should not be confused with the one-letter amino acid abbreviations on the x-axis.

Figure 4.6C shows a highly-resolved probability map generated using the 100 replica simulation while the 10 and 1

replica series shows sparsely-distributed probabilities with little agreement between the neighbouring residues. So for simplicity, only the 100 replica series will be discussed. Secondary structure 3_{10} helix has very low probability of occurrence during the simulation irrespective of amino acid residue. Other structures, like α , π helices and β -bridges also exist at low probabilities for most residues. Low probabilities can be observed for β -sheets while most residues are involved in turns, bends and loop type structures.

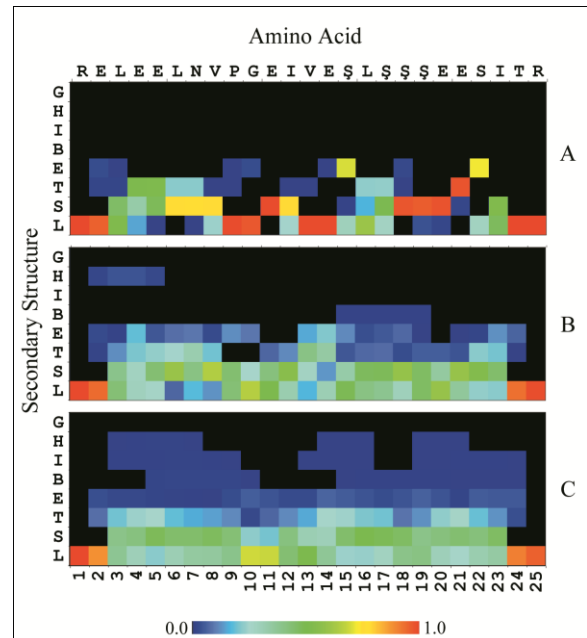


Figure 4.6 Secondary structure (SS) probability plot for (A) 1 replica, (B) 10 replicas, and (C) 100 replicas series of β -CPP simulated for a cumulative time of 1.0 μ s. (G: 3_{10} , H: α , I: π) helix, B: β -bridge, E: Extended OR β -sheet, T: Turn, S: Bend, L: Loop. Probabilities were calculated independently for each residue and a corresponding SS and are not affected by probability of adjacent residues to conform to a particular type of SS. Colour range shows probability and black regions show SS of <0.001 probability.

4.4.5.1 Helix:

There is very low probability of forming α and π helices in the various amino acid residues. Helices require turns as a prerequisite. Although all amino acids, excepting termini, are at some point involved in turns, these turns do not necessarily become helices. The main reason for this can be associated with the high degree of charge along the polypeptide chain. Formation of a helix results in a reduced R_g of the involved residues¹⁴⁸, which is not possible in the case of β -CPP due to electrostatic repulsions. This also explains why more residues are involved in π -helix which requires a longer 5-turn prerequisite and therefore would have a larger R_g of the involved residues. Larger R_g , would lower the intramolecular electrostatic repulsion. Experimental evidence for helical structural comes from three different studies. The first two studies used CD as well as many sequence-based prediction algorithms based on globular protein analysis to identify helical structures in this sequence^{68, 131}. The identified residues involved in helical structure were E₂-E₅ and S₁₉-R₂₅. The third study, conducted by Chaplin et al. in 1988¹¹⁷, isolated the fragments involved and used CD to detect helix formation. Although, in aqueous solutions, helices were not detected, upon addition of >40% 2,2,2-trifluoroethanol, helices were observed in the spectra¹¹⁷ for both regions. So, although the potential existed, the probability was low. This correlates very well to the secondary structure probability plot where E₂-E₅ and S₁₉-R₂₅ regions are involved in α and π helices. Both regions also exhibit low probabilities of helix formation.

4.4.5.2 β -Structure (Bridges, Ladders, and Sheets)

There are four sets of residues that appear to participate in β -sheet formation with moderate probability (G₁₀, E₁₄, S₁₉, and S₂₂). The low probability of bridges, which precede β -sheets, is however, surprising. This may be due to the fact that once a bridge is

formed it initiates the formation of the sheet, which would be stable for a longer period of time than a singular bridge as a result of multiple H-bonds.

4.4.5.3 Loops:

Loops are conformationally the most mobile secondary structures due to the lack of H-bonds and any restricted angle constraints. The N and C terminus arginine residues exist as loops because every secondary structure calculation is based on angles and hydrogen bonding patterns of flanking residues. This is an inherent property of terminal residues¹⁴⁴. Furthermore, these termini have been associated with loops in NMR analysis of the β -CPP peptide²⁸. Glycine (G₁₀) and glutamic acid (E₁₁) are also associated with loop-type structures with highest probability. The small size of glycine (which allows greater conformational flexibility) could be responsible for its presence in a loop. By association with a conformationally flexible neighbour, glutamic acid is also part of loop type structure. S₁₈ and S₁₉ also have a high probability of existing in loop-type structures.

4.4.5.4 Bends:

The highest likelihood for all residues within the 3-24 window is their presence in bend-type secondary structures. The absence of characterization for the first and second residues from the N and C termini is due to the inherent property of bends that requires the two amino acids on each side of the residue of interest, in order to perform the calculation. The classification criterion of turns is the presence of a hydrogen bond between C=O (*i*) and N-H (*i*+*n*), where *i* is the residue location and *n*=2,3,4,5 (γ , β , α , π). In addition, a 3-turn can take up various other dihedrals which have very high rotation angles¹³⁸. A bend on the other hand is classified simply by the sum of three backbone rotation angles $>70^\circ$ and a C ^{α} (*i*) to C ^{α} (*i*+3) distance of $<7\text{\AA}$ ¹³⁸. The difference between a turn and a bend is in the presence or absence of a hydrogen bond, respectively. So it is

not surprising that in regions with a high likelihood of turns, there is also the presence of bends (Figure 4.6).

4.4.5.5 Turns:

As predicted from the H-bond probability map, there are three clusters of residues that can exist as turns; E₂-P₉, E₁₁-S₁₈ and S₁₉-T₂₄. This does not indicate the connectivity, but rather the range of amino acids possibly involved in turns. Based on H-bond probability plot, L₃-E₅, E₁₁-V₁₃, S₁₈, V₁₄-L₁₆ and E₂₀-S₂₂ appear to be involved in turns. The lack of turn probability at glycine G₁₀ is in agreement with literature simulations on proline and glycine-rich peptides¹⁴⁹. It has been found that glycine content is anti-correlated to turn formation in peptides¹⁴⁹. This is because high glycine content increases the entropic cost of constraining the backbone in the formation of turns¹⁴⁹. Naturally, this is aggravated by proline (P₉) that neighbours the glycine (G₁₀) residue on β -CPP and has also been found to be anti-correlated with formation of turns and β -sheets. This anti-correlation with turns has the effect of increasing the hydration of the peptide backbone due to the allowable conformational movement¹⁴⁹.

Experimentally, only the V₈-E₁₁ turn has been consistently predicted from Chou-Fasman algorithms^{132, 133} as well as NMR¹⁵⁰. Although the potential of forming a turn is there, the probability of glycine (G₁₀) existing in a turn-type structure is greatly reduced. This is contradictory to the conclusions set forth by Cross *et al* (2005)²⁸, but in line with the conclusions of Tsuda *et al* in 1991¹¹⁸. In both studies, the d_α chemical shift between P₉ to E₁₁ (which would have ensured the presence of a turn sequence), was absent. However, the neighbouring residues had the potential to form turns. So, although the results were similar, the authors' conclusions were different. The reason for Cross *et al.* (2001) to

classify this region as a turn was because of other peaks associated with P₉ and E₁₁ which suggested distance of less than 0.5nm. This does not necessarily mean that there is a H-bond mediated stabilization when it could also be a bend type structure. In addition Cross et al. (2001) also suggested that turn regions included S₁₇-E₂₀, and E₂₁-T₂₄. The simulation suggests otherwise and coincidentally, relative to turn probabilities, both of these regions are consistently associated with high probabilities of bends and loop-type structures (Figure 4.6), which lends to the difficulty in associating structural features with NMR chemical shifts.

4.4.6 Ramachandran Angles

The DSSP probability plot shows the existence of secondary structure. However, one structure that is not classified, but has been known to exist in β -CPP, is the PPII structure¹¹⁶. Structural characterization of

PPII structure is based on dihedral angles in the region around $\phi: -60^\circ$, $\psi: 140^\circ$. Recall that plotting ϕ and ψ angles results in a distribution plot which can indicate propensity of angles. This can be correlated to various types of secondary structures. Simulation of phosphorylated serine groups indicates presence of some torsion angles to the left of the helix region (G, H and I in Figure 4.7).

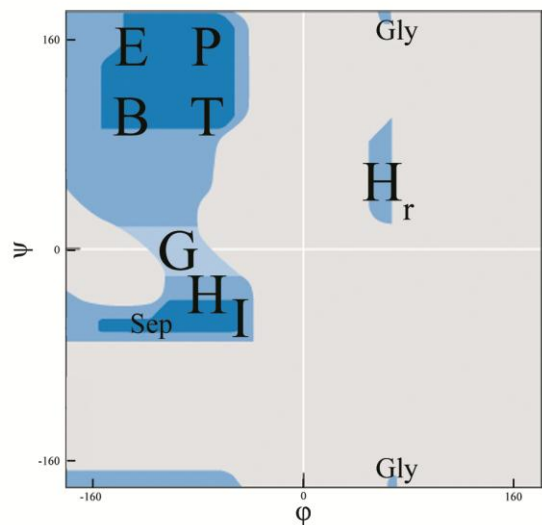


Figure 4.7 Theoretical divisions of the Ramachandran plot according to various secondary structures as well as some special angles for amino acids. G: 3₁₀, H: α , I: π) helix, B: β -bridge, E: Extended OR β -sheet, T: Turn, P: PPII Helix, H_r: Right-Handed Helix, Gly: Glycine, and Sep: Phosphoserine.¹⁵¹⁻¹⁵³

The Ramachandran plot shows the probability distribution of dihedral angles through the trajectory using the rainbow-coloured lookup table where red indicates favourable regions and blue indicates unfavourable dihedrals (Figure 4.8). The plot shows that the most allowed dihedral angles in the area of $\phi < -20^\circ, \psi > -80^\circ$, placing them in the region of α -helices and β -sheets and some very low lying probabilities for $\phi > -20^\circ$. The region for β -sheets was in the area confined by $\phi < -20^\circ, \psi > 80^\circ$ and exhibited a bilobal shape at $\sim -95^\circ > \phi > -100^\circ$. On the other hand, the region for α -helices enclosed by the angles $\phi < -20^\circ, 0^\circ > \psi > -80^\circ$ was separated into two sections with both areas equally populated. In addition, the angles $\phi: -120^\circ$ and $\phi - 70^\circ$ in the range $140 < \psi < 145^\circ$ shown in dark red indicated the highest dihedral angles in the peptide.

Although high frequencies of these regions are implicated in the presence of these two different secondary structures, accounting the secondary

structure analysis (Figure 4.6) shows that these angles are indicative only of the propensity of the 23 dihedral angles in the peptide. This means that there are forces that

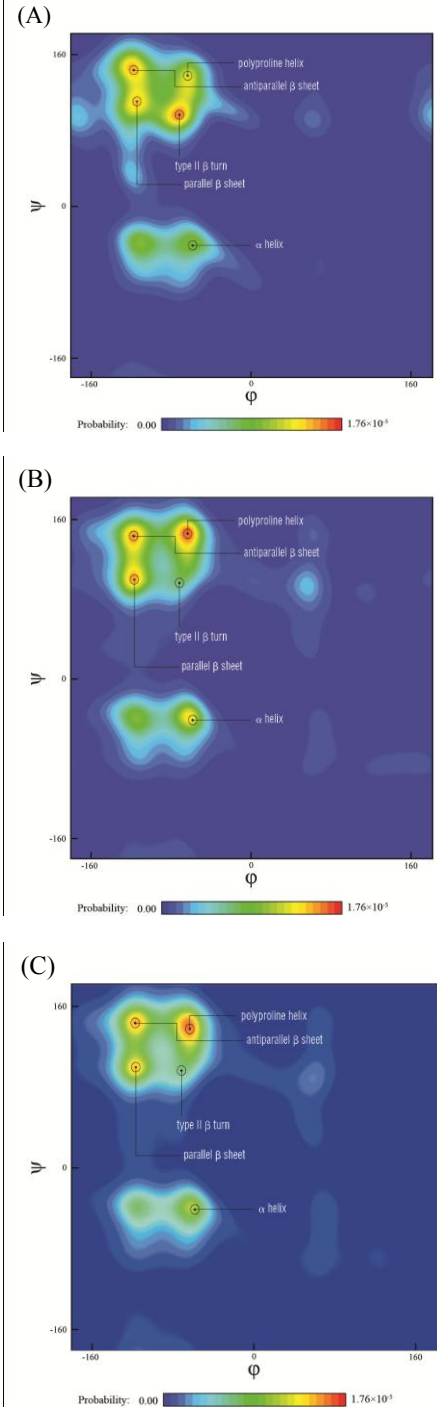


Figure 4.8 Ramachandran Angles for (A) 1,(B) 10,(C) 100 replica series.

may prevent peptidic domains from folding into the appropriate secondary structures. This is not surprising as there are a number of Glutamic acid residues in the peptide composition, which is the second most structure-dissolving amino acid in the DisProt database²¹. In addition, this is aggravated by the presence of proline whose cyclic structure rigidifies the ϕ portion of its dihedral preventing formation of secondary structure elements¹⁵⁴. Although β -CPP contains only one proline, the highest region frequented in the Ramachandran plot is the region around $\phi: -60^\circ$, $\psi: 140^\circ$. This region is associated with the PPII structure, which exists in certain peptides even when there is a paucity of prolines. With higher number of replicas, the only main difference in the plots is that the population of angles corresponding to PPII helices are increased. Another important set of Ramachandran angles observed are the angles to the left of the helix region. This region has been observed specifically for phosphoserine residues in simulations^{152, 153}.

Interestingly, increasing the replica number did not have a significant impact on the contours of the Ramachandran plot as compared to H-bonds and DSSP probability plot. With even one long simulation, it was possible to determine the contours of the dihedral angles in the plot. This has two important implications for simulation methodology. First, Ramachandran plots cannot be used to as quality control measures or to determine the performance of simulation as the angles populated with 1, 10 or a 100 replica series are generally the same. Second, these plots are useful only to determine propensity of angles, but not definite structures. Third, the local angle population of a few replicas is representative of the conformational ensemble.

Experimental cross-validation with Raman Optical Activity (ROA) of β -casein indicated two populations of dihedral angles. The more favoured population is that of the PPII structure, while the minor population was of a more extended, but hydrated, peptide bonds that mitigated the possible presence of β -sheets²². Since the ROA study was conducted with β -casein, an exact comparison is not present, although some similarities can be seen in the Ramachandran plot. The dihedral populations that appear in the highest density were of the PPII structure as per ROA analysis²². However, unlike ROA, the simulation showed a lower propensity towards α -helix type dihedrals. It is possible that the lower probability of this population of dihedrals prevented its detection in the ROA spectrum.

A qualitative NMR study on conformation of β -CPP in a variety of solutions showed that in the case where the peptide was neutralized by NaCl ions, there was a propensity towards β -turns as well as α -helical angles on the Ramachandran plot²⁸. The most recent CD evidence also suggests low level of α -helix (5%), and high levels of β -structure and turns (34, 28%)¹¹⁶. Hence, the simulation is in qualitative agreement with experimental studies conducted thus far.

4.4.7 Working Hypothesis for interaction with calcium and phosphate ions

Combination of all data for Rg, H-bond, secondary structure as well as Ramachandran plot alludes to a mechanism related to the function of β -CPP. In the presence of calcium and phosphate ions, β -CPP forms nanoclusters with very precise Rg (2.80 ± 0.05 nm) (10mg/mL with 37mM of Ca^{2+} and 30mM of PO_4^{3-} ions)¹⁴⁰. The mechanism of such a controlled Rg is currently unknown other than the fact that it is associated with the disorder of the peptide^{20, 155}. On its own, the peptide exhibits a distribution of Rg ranging

from 0.75-1.1nm with most conformations at 0.84nm. When complexed with Ca^{2+} and PO_4^{3-} ions, the terminal arginines are displaced to give way to binding to the phosphoseryl cluster. Consequently, the size of the peptide would increase when in the form of a complexed peptide, as has been observed with SAXS and SANS data¹⁵⁵. As a result, these terminal arginines become involved with residues prior to the phosphorylated motif. In addition, there is a significant amount of conformational flexibility at glycine G_{10} by virtue of its presence in a loop. Thus, the hypothesized mechanism of binding of β -CPP to Ca^{2+} is as follows: Ca^{2+} and PO_4^{3-} would bind to the phosphoseryl residues at residues 15-19. The mineralization would grow as a result of nucleation. To limit nucleation, the size of the calcium phosphate nanocluster would be mitigated by the conformational flexibility offered by glycine at G_{10} . The flexibility would be exhibited by the N-terminus arm of the peptide, which would increase local entropy and prevent binding. The N-terminus arginine would bind electrostatically with amino acids prior to the phosphorylated motif, which would limit binding of Ca^{2+} and PO_4^{3-} ions to only the phosphoseryl residues. This would allow the peptide to stabilize the colloid in aqueous solutions. As this is a working hypothesis, in order to obtain a complete picture there are two important future studies that must be conducted. First, a simulation of the peptide in presence of calcium and phosphate ions to characterize such structural features must be conducted. A second simulation of multiple peptides in the presence of calcium and phosphate ions to determine the structure of the calcium phosphate nanocluster bound to β -CPPs is also necessary.

4.5 Conclusions

A molecular dynamics simulation method to determine casein structure was devised. The method was tested with β -CPP and yielded qualitative agreement with existing experiment results. Radius of gyration ranged from 0.75-1.1nm. H-bond analysis indicated highest probability for local bonds 3 residues apart rather than longer range interactions. The secondary structure is in agreement with experimental as well as simulation studies conducted so far. Ramachandran plot showed four major populations of dihedral angles including PPII angles as well as angles from phosphoserine residues, in qualitative agreement with ROA, NMR, and simulation data. Most individual amino acids have preference towards bends which precede turns, except for glycine. The conformational flexibility of glycine allows for a loop-type structure. It is hypothesized that the conformational flexibility of glycine prevents mineral growth after sequestration due to local entropy. This can be tested with future simulations of the peptide in presence of calcium and phosphate ions.

4.5.1 Future Directions

Knowing the conformation ensemble of β -CPP can allow understanding of calcium phosphate sequestration phenomena as well as how the peptide can increase the bioavailability of calcium phosphate. As NMR spectral changes have been observed in the polypeptide upon addition of calcium ions, conformational studies of β -CPP under different physicochemical conditions could help identify the exact mechanism of controlling the size of the amorphous calcium phosphate nanocluster. As the peptide is involved with increasing the bioavailability of calcium, simulation studies with lipid bilayers could provide atomic scale resolution on the phenomena. Another important

aspect that could be studied in detail is atomic scale interaction of the Fab region of IgE. This could identify structural components required for recognition. Finally, since an MD method has been developed and validated with experimental results, it is possible to test the method with other important regions of the casein micelle which may be involved with VD₃ binding – the original goal of the thesis.

4.6 Acknowledgements

The authors recognize funding support from AFMNet, Dairy Farmers of Canada, and NSERC. Special thanks to Dr. David Pink, Jordan Marsh, Dr. Shajahan Gulam Razul from St. Francis Xavier University and Daniel Gruner from Scinet for their help in setting up the software to run the simulations. Dr. Carl Holt is thanked for his insight on casein secondary structure determination and Dr. Harold Farrell is thanked for important discussions about Ramachandran plot as well as the β -CPP PDB file from his simulations. Dr. Régis Pomès, his Ph.D. student Sarah Rauscher and their lab group at the Hospital for Sick Children are acknowledged for their expertise with simulations and for guiding with the simulation methodology and continued technical support. Justin Lemkul is acknowledged for providing the modified Gromos 43a1p files the parameters of which were calculated by Graham Smith. Ace-net and Scinet are sincerely acknowledged for their computational resources, without which this study would have taken eons to complete.

Chapter 5: Overall Remarks

The development of a simple binding assay now affords the ability to study interactions between macromolecules and vitamin D₃. In terms of the initial goal of understanding the interaction of vitamin D₃ with casein, this was an important milestone. With future optimization of VD₃ – casein binding interaction as well as the reliable carryover of VD₃ from milk to cheese, it is easy to foresee that the technology of VD₃ fortified cheese using casein will come to fruition within 5-10 years.

Regarding the molecular understanding of casein and VD₃, the end appears far. However, by performing simulations of a single casein peptide, the stage has been set for future determination of β -casein and casein micelle structure. Simplification of the simulation problems may be required to obtain the statistical ensemble quickly and efficiently. Once the structure of casein micelle is understood, docking simulations with VD₃ would bridge the gap in understanding between binding interaction and localization of VD₃ inside the casein micelle. Overall, although there is a lot to be done in order to understand the interaction of casein with VD₃, what this thesis has done is lay the foundation.

References

1. Jones, G.; Strugnelli, S. A.; DeLuca, H. F., *Physiological Reviews* **1998**, 78 (4), 1193-1231.
2. Holick, M. F., *New England Journal of Medicine* **2007**, 357 (3), 266-281.
3. Vieth, R.; Bischoff-Ferrari, H.; Boucher, B. J.; Dawson-Hughes, B.; Garland, C. F.; Heaney, R. P.; Holick, M. F.; Hollis, B. W.; Lamberg-Allardt, C.; McGrath, J. J.; Norman, A. W.; Scragg, R.; Whiting, S. J.; Willett, W. C.; Zittermann, A., *American Journal of Clinical Nutrition* **2007**, 85 (3), 649-650.
4. DeLuca, H. F.; Zierold, C., *Nutrition Reviews* **1998**, 56 (2), S4-S10.
5. Calvo, M. S.; Whiting, S. J.; Barton, C. N., *American Journal of Clinical Nutrition* **2004**, 80 (6), 1710S-1716S.
6. Banville, C.; Vuillemand, J. C.; Lacroix, C., *International Dairy Journal* **2000**, 10 (5-6), 375-382.
7. Kazmi, S. A.; Vieth, R.; Rousseau, D., *International Dairy Journal* **2007**, 17 (7), 753-759.
8. Wagner, D.; Sidhom, G.; Whiting, S. J.; Rousseau, D.; Vieth, R., *Journal of Nutrition* **2008**, 138 (7), 1365-1371.
9. Langlois, K.; Greene-Finestone, L.; Little, J.; Hidirolou, N.; Whiting, S., *Health Rep* **2010**, 21 (1), 47-55.
10. Schmit, T. M.; Kaiser, H. M., *Journal of Dairy Science* **2006**, 89 (12), 4924-4936.
11. Yang, M. C.; Guan, H. H.; Liu, M. Y.; Lin, Y. H.; Yang, J. M.; Chen, W. L.; Chen, C. J.; Mao, S. J. T., *Proteins-Structure Function and Bioinformatics* **2008**, 71 (3), 1197-1210.
12. Yang, M. C.; Chen, N. C.; Chen, C. J.; Wu, C. Y.; Mao, S. J. T., *Febs Journal* **2009**, 276 (8), 2251-2265.
13. Lucey, J. A., *Journal of Dairy Science* **2002**, 85 (2), 281-294.
14. Semo, E.; Kesselman, E.; Danino, D.; Livney, Y. D., *Food Hydrocolloids* **2007**, 21 (5-6), 936-942.
15. Horne, D. S., *Current Opinion in Colloid & Interface Science* **2006**, 11 (2-3), 148-153.
16. Farrell, H. M.; Jimenez-Flores, R.; Bleck, G. T.; Brown, E. M.; Butler, J. E.; Creamer, L. K.; Hicks, C. L.; Hollar, C. M.; Ng-Kwai-Hang, K. F.; Swaisgood, H. E., *Journal of Dairy Science* **2004**, 87 (6), 1641-1674.
17. Tuinier, R.; de Kruif, C. G., *Journal of Chemical Physics* **2002**, 117 (3), 1290-1295.
18. Holt, C.; Sawyer, L., *Journal of the Chemical Society-Faraday Transactions* **1993**, 89 (15), 2683-2692.
19. Smyth, E.; Clegg, R. A.; Holt, C. In *A biological perspective on the structure and function of caseins and casein micelles*, 2004; pp 121-126.
20. Little, E. M.; Holt, C., *European Biophysics Journal with Biophysics Letters* **2004**, 33 (5), 435-447.
21. Rauscher, S.; Pomes, R., *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire* **2010**, 88 (2), 269-290.

22. Syme, C. D.; Blanch, E. W.; Holt, C.; Jakes, R.; Goedert, M.; Hecht, L.; Barron, L. D., *European Journal of Biochemistry* **2002**, 269 (1), 148-156.
23. Holt, C.; Sawyer, L., *Protein Engineering* **1987**, 1 (3), 243-243.
24. Horne, D. S., *Current Opinion in Colloid & Interface Science* **2002**, 7 (5-6), 456-461.
25. Berendsen, H. J. C., *Science* **1998**, 282 (5389), 642-643.
26. Rauscher, S.; Pomès, R., *Journal of Physics: Conference Series* **2010**, 256, 1-13.
27. Holt, C.; Sorensen, E. S.; Clegg, R. A., *Febs Journal* **2009**, 276 (8), 2308-2323.
28. Cross, K. J.; Huq, N. L.; Bicknell, W.; Reynolds, E. C., *Biochemical Journal* **2001**, 356, 277-285.
29. Kumosinski, T. F.; Brown, E. M.; Farrell, H. M., *Journal of Dairy Science* **1993**, 76 (4), 931-945.
30. Kumosinski, T. F.; Brown, E. M.; Farrell, H. M., *Journal of Dairy Science* **1991**, 74 (9), 2889-2895.
31. Kumosinski, T. F.; Brown, E. M.; Farrell, H. M., *Journal of Dairy Science* **1991**, 74 (9), 2879-2889.
32. Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M., *Journal of the American Chemical Society* **2005**, 127 (2), 476-477.
33. Holick, M. F., *American Journal of Clinical Nutrition* **2004**, 79 (3), 362-371.
34. Lin, R.; White, J. H., *Bioessays* **2004**, 26 (1), 21-28.
35. Haug, A.; Hostmark, A. T.; Harstad, O. M., *Lipids in Health and Disease* **2007**, 6.
36. Milne, G. W. A.; Delander, M., *Vitamin D handbook : structures, synonyms, and properties*. Wiley-Interscience: Hoboken, N.J., 2008; p ix, 274 p.
37. Olson, R. E., *Journal of Nutrition* **1998**, 128 (2), 439S-443S.
38. Vieth, R.; Cole, D. E.; Hawker, G. A.; Trang, H. M.; Rubin, L. A., *European Journal of Clinical Nutrition* **2001**, 55 (12), 1091-1097.
39. Holick, M. F., *Annals of the New York Academy of Sciences* **1988**, 548, 14-26.
40. Holick, M. F.; Chen, T. C., *American Journal of Clinical Nutrition* **2008**, 87 (4), 1080S-1086S.
41. Ross, A. C.; Institute of Medicine (U.S.). Committee to Review Dietary Reference Intakes for Vitamin D and Calcium., *DRI, dietary reference intakes : calcium, vitamin D*. National Academies Press: Washington, D.C., 2011; p xv, 1115 p.
42. Dawson-Hughes, B.; Heaney, R. P.; Holick, M. F.; Lips, P.; Meunier, P. J.; Vieth, R., *Osteoporosis International* **2005**, 16 (7), 713-716.
43. Faulkner, H.; Hussein, A.; Foran, M.; Szijarto, L., *Journal of Dairy Science* **2000**, 83 (6), 1210-1216.
44. Carroccio, A.; Montalto, G.; Cavera, G.; Notarbatolo, A.; Lactase Deficiency Study, G., *Journal of the American College of Nutrition* **1998**, 17 (6), 631-636.
45. Lovelace, H. Y.; Barr, S. I., *Journal of the American College of Nutrition* **2005**, 24 (1), 51-57.
46. Sherman, P. M., *Canadian Journal of Gastroenterology* **2004**, 18 (2), 81-82.
47. USDA National Nutrient Database for Standard Reference, Release 23. 2010 ed.; U.S. Department of Agriculture, Agricultural Research Service: 2010.
48. Yetley, E. A., *American Journal of Clinical Nutrition* **2008**, 88 (2), 558S-564S.
49. Dalgleish, D. G., *Soft Matter* **2011**, 7 (6), 2265-2272.
50. Thompson, A.; Boland, M.; Singh, H.; Knovel (Firm), Milk proteins

from expression to food. In *Food science and technology international series* [Online] Academic Press/Elsevier,; Amsterdam ; Boston, 2009; pp. xv, 535 p., 8 p. of plates. <http://www.knovel.com/knovel2/Toc.jsp?BookID=3003>.

51. Forrest, S. A.; Yada, R. Y.; Rousseau, D., *Journal of Agricultural and Food Chemistry* **2005**, 53 (20), 8003-8009.
52. Wang, Q. W.; Allen, J. C.; Swaisgood, H. E., *Journal of Dairy Science* **1997**, 80 (6), 1054-1059.
53. Flower, D. R., *Biochemical Journal* **1996**, 318, 1-14.
54. Wu, S. Y.; Perez, M. D.; Puyol, P.; Sawyer, L., *Journal of Biological Chemistry* **1999**, 274 (1), 170-174.
55. Qin, B. Y.; Bewley, M. C.; Creamer, L. K.; Baker, H. M.; Baker, E. N.; Jameson, G. B., *Biochemistry* **1998**, 37 (40), 14014-14023.
56. Taulier, N.; Chalikian, T. V., *Journal of Molecular Biology* **2001**, 314 (4), 873-889.
57. Zimet, P.; Livney, Y. D., *Food Hydrocolloids* **2009**, 23 (4), 1120-1126.
58. Liang, L.; Tajmir-Riahi, H. A.; Subirade, M., *Biomacromolecules* **2008**, 9 (1), 50-56.
59. Considine, T.; Patel, H. A.; Singh, H.; Creamer, L. K., *Food Chemistry* **2007**, 102 (4), 1270-1280.
60. Holt, C.; Sawyer, L., *Journal of the Chemical Society-Faraday Transactions* **1993**, 89 (15), 2683-2692.
61. Ginger, M. R.; Grigor, M. R., *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* **1999**, 124 (2), 133-145.
62. Home, D. S., *Current Opinion in Colloid & Interface Science* **2006**, 11 (2-3), 148-153.
63. Horne, D. S., *Current Opinion in Colloid & Interface Science* **2002**, 7 (5-6), 456-461.
64. Fox, P. F.; Brodtkorb, A., *International Dairy Journal* **2008**, 18 (7), 677-684.
65. Dalgleish, D. G., *Lait* **2007**, 87 (4-5), 385-387.
66. de Kruif, C. G. In *Casein micelle interactions*, 1999; pp 183-188.
67. Walstra, P. In *Casein sub-micelles: do they exist?*, 1999; pp 189-192.
68. Sawyer, L.; Holt, C., *Journal of Dairy Science* **1993**, 76 (10), 3062-3078.
69. Wright, P. E.; Dyson, H. J., *Journal of Molecular Biology* **1999**, 293 (2), 321-331.
70. Tompa, P., *Trends in Biochemical Sciences* **2002**, 27 (10), 527-533.
71. Leermakers, F. A. M.; Atkinson, P. J.; Dickinson, E.; Horne, D. S., *Journal of Colloid and Interface Science* **1996**, 178 (2), 681-693.
72. Dickinson, E., *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **2006**, 288 (1-3), 3-11.
73. Holt, C., *International Dairy Journal* **1991**, 1, 151-165.
74. Leclerc, E.; Calmettes, P., *Physical Review Letters* **1997**, 78 (1), 150-153.
75. Pink, D. A.; Quinn, B.; Baskin, K., *Langmuir* **1994**, 10 (8), 2559-2565.
76. Thomason, P.; Kay, R., *Journal of Cell Science* **2000**, 113 (18), 3141-3150.
77. Chen, Y.; Bal, B. S.; Gorski, J. P., *Journal of Biological Chemistry* **1992**, 267 (34), 24871-24878.
78. Kawasaki, K.; Weiss, K. M., *Proceedings of the National Academy of Sciences of the United States of America* **2003**, 100 (7), 4060-4065.

79. Marsh, M. E.; Sass, R. L., *Journal of Experimental Zoology* **1983**, 226 (2), 193-203.
80. Huppertz, T.; de Kruif, C. G., *International Dairy Journal* **2008**, 18 (5), 556-565.
81. Pitkowski, A.; Nicolai, T.; Durand, D., *Biomacromolecules* **2008**, 9 (1), 369-375.
82. Huppertz, T.; Vaia, B.; Smiddy, M. A., *Journal of Dairy Research* **2008**, 75 (1), 44-47.
83. Cross, K. J.; Huq, N. L.; Reynolds, E. C., *Current Pharmaceutical Design* **2007**, 13 (8), 793-800.
84. Cross, K. J.; Huq, N. L.; Palamara, J. E.; Perich, J. W.; Reynolds, E. C., *Journal of Biological Chemistry* **2005**, 280 (15), 15362-15369.
85. Chatchatee, P.; Jarvinen, K. M.; Bardina, L.; Vila, L.; Beyer, K.; Sampson, H. A., *Clinical and Experimental Allergy* **2001**, 31 (8), 1256-1262.
86. Sampson, H. A.; Mendelson, L.; Rosen, J. P., *New England Journal of Medicine* **1992**, 327 (6), 380-384.
87. Calvo, M. S.; Whiting, S. J.; Barton, C. N., *Journal of Nutrition* **2005**, 135 (2), 310-316.
88. van de Weert, M., *Journal of Fluorescence* **2010**, 20 (2), 625-629.
89. Bouchemal, K.; Agnely, F.; Koffi, A.; Djabourov, M.; Ponchel, G., *Journal of Molecular Recognition* **2010**, 23 (4), 335-342.
90. Salim, N. N.; Feig, A. L., *Methods* **2009**, 47 (3), 198-205.
91. Heerklottz, H. H.; Binder, H.; Epand, R. M., *Biophysical Journal* **1999**, 76 (5), 2606-2613.
92. Bouchemal, K., *Drug Discovery Today* **2008**, 13 (21-22), 960-972.
93. Dai, W.-G.; Pollock-Dove, C.; Dong, L. C.; Li, S., *Advanced Drug Delivery Reviews* **2008**, 60 (6), 657-672.
94. Ganesan, B.; Brothersen, C.; McMahon, D. J., *Journal of Dairy Science* **2011**, 94 (7), 3708-3714.
95. Narayan, M.; Berliner, L. J., *Protein Science* **1998**, 7 (1), 150-157.
96. Spector, A. A.; Fletcher, J. E., *Lipids* **1970**, 5 (4), 403-&.
97. Corredig, M.; Dalgleish, D. G., *Journal of Dairy Research* **1996**, 63 (3), 441-449.
98. Hayes, M. G.; Fox, P. F.; Kelly, A. L., *Journal of Dairy Research* **2005**, 72 (1), 25-33.
99. Pitha, J.; Milecki, J.; Fales, H.; Pannell, L.; Uekama, K., *International Journal of Pharmaceutics* **1986**, 29 (1), 73-82.
100. Freifelder, D., *Physical biochemistry : applications to biochemistry and molecular biology*. 2nd ed.; W.H. Freeman: San Francisco, 1982; p xii, 761 p.
101. Nitkowski, A.; Baeumner, A.; Lipson, M.-c. s. f. b. u. m. r., *Biomedical Optics Express* **2011**, 2 (2), 6.
102. Loftsson, T.; Hreinsdottir, D., *Aaps Pharmscitech* **2006**, 7 (1).
103. Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V.; Radchenko, E.; Zefirov, N. S.; Makarenko, A. S.; Tanchuk, V. Y.; Prokopenko, V. V., *Journal of Computer-Aided Molecular Design* **2005**, 19 (6), 453-463.
104. Merce, A. L. R.; Nicolini, J.; Khan, M. A.; Bouet, G., *Carbohydrate Polymers* **2009**, 77 (2), 402-409.
105. Portnaya, I.; Cogan, U.; Livney, Y. D.; Ramon, O.; Shimoni, K.; Rosenberg, M.; Danino, D., *Journal of Agricultural and Food Chemistry* **2006**, 54 (15), 5555-5561.

106. Anema, S. G.; Li, Y. M., *Journal of Dairy Research* **2003**, 70 (1), 73-83.
107. Faka, M.; Lewis, M. J.; Grandison, A. S.; Deeth, H., *International Dairy Journal* **2009**, 19 (6-7), 386-392.
108. Berry, G. P.; Creamer, L. K., *Biochemistry* **1975**, 14 (16), 3542-3545.
109. Holt, C.; Wahlgren, N. M.; Drakenberg, T., *Biochemical Journal* **1996**, 314, 1035-1039.
110. Wang, X.; Megert, B.; Hellwig, E.; Neuhaus, K. W.; Lussi, A., *Journal of Dentistry* **2011**, 39 (2), 163-170.
111. Rahiotis, C.; Vougiouklakis, G.; Eliades, G., *Journal of Dentistry* **2008**, 36 (4), 272-280.
112. Ferraretto, A.; Gravaghi, C.; Fiorilli, A.; Tettamanti, G., *Febs Letters* **2003**, 551 (1-3), 92-98.
113. Boutrou, R.; Coirre, E.; Jardin, J.; Leonil, J., *Journal of Agricultural and Food Chemistry* **2010**, 58 (13), 7955-7961.
114. Otani, H.; Mine, Y.; Hosono, A., *Milchwissenschaft-Milk Science International* **1987**, 42 (8), 505-508.
115. Mizumachi, K.; Kurisaki, J.; Kaminogawa, S., *Bioscience Biotechnology and Biochemistry* **1999**, 63 (5), 911-916.
116. Farrell, H. M.; Qi, P. X.; Wickham, E. D.; Unruh, J. J., *Journal of Protein Chemistry* **2002**, 21 (5), 307-321.
117. Chaplin, L. C.; Clark, D. C.; Smith, L. J., *Biochimica Et Biophysica Acta* **1988**, 956 (2), 162-172.
118. Tsuda, S.; Niki, R.; Kuwata, T.; Tanaka, I.; Hikichi, K., *Magnetic Resonance in Chemistry* **1991**, 29 (11), 1097-1102.
119. Wahlgren, N. M.; Leonil, J.; Dejmek, P.; Drakenberg, T., *Biochimica Et Biophysica Acta* **1993**, 1202 (1), 121-128.
120. Farrell, H. M.; Wickham, E. D.; Unruh, J. J.; Qi, P. X.; Hoagland, P. D., *Food Hydrocolloids* **2001**, 15 (4-6), 341-354.
121. Curley, D. M.; Kumosinski, T. F.; Unruh, J. J.; Farrell, H. M., *Journal of Dairy Science* **1998**, 81 (12), 3154-3162.
122. Byler, D. M.; Farrell, H. M.; Susi, H., *Journal of Dairy Science* **1988**, 71 (10), 2622-2629.
123. Kumosinski, T. F.; Brown, E. M.; Farrell, H. M., *Journal of Dairy Science* **1993**, 76 (4), 931-945.
124. Yasar, F.; Celik, S.; Koksel, H., *Physica a-Statistical Mechanics and Its Applications* **2006**, 363 (2), 348-358.
125. Chou, P. Y.; G.D., F., *Advanced Enzymology* **1978**, 4.
126. Garnier, J.; Osguthorpe, D. J.; Robson, B., *Journal of Molecular Biology* **1978**, 120 (1), 97-120.
127. Robson, B.; Pain, R. H., *Journal of Molecular Biology* **1971**, 58 (1), 237-&.
128. Schiffer, M.; Edmundso.Ab, *Biophysical Journal* **1967**, 7 (2), 121-&.
129. Lim, V. I., *Journal of Molecular Biology* **1974**, 88 (4), 873-894.
130. Kabat, E. A.; Wu, T. T., *Annals of the New York Academy of Sciences* **1971**, 190, 382-&.
131. Creamer, L. K.; Richardson, T.; Parry, D. A. D., *Archives of Biochemistry and Biophysics* **1981**, 211 (2), 689-696.

132. Bloomfield, V. A.; Mead, R. J., *Journal of Dairy Science* **1975**, 58 (4), 592-601.
133. Graham, E. R. B.; Malcolm, G. N.; McKenzie, H. A., *International Journal of Biological Macromolecules* **1984**, 6 (3), 155-161.
134. Cassiano, M. M.; Areas, J. A. G., *Journal of Dairy Science* **2003**, 86 (12), 3876-3880.
135. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., *Journal of Chemical Theory and Computation* **2008**, 4 (3), 435-447.
136. Wu, W.; Kerrigan, J. E.; Yadav, P.; Schwartz, B.; Izotova, L.; Lavoie, T. B.; Pestka, S., *Oncology Research* **2004**, 14 (11-12), 541-558.
137. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., *Journal of Chemical Physics* **1984**, 81 (8), 3684-3690.
138. Kabsch, W.; Sander, C., *Biopolymers* **1983**, 22 (12), 2577-2637.
139. Narang, P.; Bhushan, K.; Bose, S.; Jayaram, B., *Physical Chemistry Chemical Physics* **2005**, 7 (11), 2364-2375.
140. Holt, C.; Timmins, P. A.; Errington, N.; Leaver, J., *European Journal of Biochemistry* **1998**, 252 (1), 73-78.
141. Liu, Y.; Guo, R., *Journal of Colloid and Interface Science* **2007**, 315 (2), 685-692.
142. Hansen, J. L.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A., *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99 (18), 11670-11675.
143. Purves, W. K., *Life, the science of biology*. 7th ed.; Sinauer Associates ; W.H. Freeman and Co.: Sunderland, Mass. Gordonsville, VA, 2004.
144. Carter, P.; Andersen, C. A. F.; Rost, B., *Nucleic Acids Research* **2003**, 31 (13), 3293-3295.
145. Marcelino, A. M. C.; Gierasch, L. M., *Biopolymers* **2008**, 89 (5), 380-391.
146. Rose, G. D.; Gierasch, L. M.; Smith, J. A., *Advances in Protein Chemistry* **1985**, 37, 74-84.
147. Small, D.; Chou, P. Y.; Fasman, G. D., *Biochemical and Biophysical Research Communications* **1977**, 79 (1), 341-346.
148. Dumont, C.; Matsumura, Y.; Kim, S. J.; Li, J.; Kondrashkina, E.; Kihara, H.; Gruebele, M., *Protein Science* **2006**, 15 (11), 2596-2604.
149. Rauscher, S.; Baud, S.; Miao, M.; Keeley, F. W.; Pomes, R., *Structure* **2006**, 14 (11), 1667-1676.
150. Cross, K. J.; Huq, N. L.; Stanton, D. P.; Sum, M.; Reynolds, E. C., *Biomaterials* **2004**, 25 (20), 5061-5069.
151. Shu, N.; Hovmoller, S.; Zhou, T., *Current Protein & Peptide Science* **2008**, 9 (4), 310-324.
152. Shen, T. Y.; Zong, C. H.; Hamelberg, D.; McCammon, J. A.; Wolynes, P. G., *Faseb Journal* **2005**, 19 (11), 1389-1395.
153. Shen, T. Y.; Wong, C. F.; McCammon, J. A., *Journal of the American Chemical Society* **2001**, 123 (37), 9107-9111.
154. Jiang, P.; Xu, W. X.; Mu, Y. G., *Plos Computational Biology* **2009**, 5 (4), 13.

155. Holt, C., *European Biophysics Journal with Biophysics Letters* **2004**, 33 (5), 421-434.