International Dairy Journal 46 (2015) 2–11

Contents lists available at ScienceDirect

International Dairy Journal

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

Review

Casein structures in the context of unfolded proteins

David C. Thorn [a], Heath Ecroyd [b], John A. Carver [c], Carl Holt [d, *]

[a] Centre for Protein Engineering, University of Liège, 4000 Liège, Belgium
[b] School of Biological Sciences, Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong, NSW 2522, Australia
[c] Research School of Chemistry, College of Physical and Mathematical Sciences, The Australian National University, Canberra, ACT 0200, Australia
[d] Institute of Molecular, Cell and Systems Biology, School of Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK

ARTICLE INFO

Article history:
Available online 30 July 2014

ABSTRACT

Caseins were among the first proteins to be recognised as functional but unfolded. Many others are now known, providing better models of casein behaviour than either detergents or folded proteins. Caseins are members of a paralogous group of unfolded phosphoproteins, some of which share the ability to sequester amorphous calcium phosphate through phosphate centres. Non-covalent interactions of caseins can be through Pro- and Gln-rich sequences. Similar sequences in other unfolded proteins can also form open and highly hydrated structures such as gels, mucus and slimes. Many unfolded proteins, including κ- and α_{S2}-caseins, can form amyloid fibrils under physiological conditions. The sequence-specific interactions that lead to fibrils can be reduced or eliminated by low specificity interactions among a mixture of caseins to yield, instead, amorphous aggregates. The size of amorphous whole casein aggregates is limited by the C-terminal half of κ-casein whose sequence resembles that of a soluble mucin.

Contents

1. Introduction

When, Arthur Kornberg made his famous remark that he had never met a dull enzyme (Kornberg, 1989), it reflected a life spent in uncovering the intricate relationships between structure and function among these largely globular proteins. Caseins were recognised to be unfolded as early as the 1950s (Halwer, 1954; Kresheck, 1965) when there were few other examples known to science. A senior figure in casein research later described casein as simply a denatured protein with only a nutritional function (McMeekin, 1970). This nutritional function of the unfolded conformation was thought to be the ease of digestion of casein by proteinases but this attracted little interest in the mainstream of

* Corresponding author. Tel.: +44 1292317615.
E-mail address: carl.holt@glasgow.ac.uk (C. Holt).

http://dx.doi.org/10.1016/j.idairyj.2014.07.008
0958-6946/

protein science. In the last two decades the number of structure–function studies on unfolded proteins has grown exponentially from just a few per annum in the 1990s to thousands per annum today so that the functional roles of unfolded proteins in biology are now legion (Dyson & Wright, 2005; Rose, 2002; Tompa, 2012; Uversky & Dunker, 2010).

In our view, casein research will benefit greatly from deep insights into the relationships between structure and function recently gained from studies on other unfolded proteins. In this review, we take some steps in this direction by considering the function of individual caseins and casein peptides in the sequestration of amorphous calcium phosphate to form thermodynamically stable nanoclusters. This ability is shared with osteopontin and probably with some other, closely related, unfolded phosphoproteins where the unfolded conformation is essential. The ability of individual caseins and casein peptides to form amyloid fibrils is shared with many other unfolded proteins and peptides. The ability of mixtures of different caseins to inhibit the growth of amyloid fibrils is an example of how promiscuous interactions of low sequence specificity can compete successfully with the highly specific interactions needed to form the fibrils.

In the past, the self-association of individual caseins and the association of casein mixtures in the absence of calcium ions or calcium phosphate have been likened to that of detergent molecules in which the hydrophobic effect (Cramer & Truhlar, 1992; Kauzmann, 1959) is assumed to provide the driving force (Horne, Lucey, & Choi, 2007; Mikheeva, Grinberg, Grinberg, Khokhlov, & de Kruif, 2003; Payens & Vreeman, 1982). Comparison with similar unfolded proteins suggests that an alternative driving force that involves main-chain-to-main-chain interactions of low sequence specificity is more likely, rather than the side-chain interactions of the hydrophobic effect. Such interactions lead naturally to open, extended and highly hydrated structures rather than the compact, anhydrous, domains of detergent micelles or the interior of globular proteins (Holt, Carver, Ecroyd, & Thorn, 2013).

In this review we will focus on two non-nutritional functions of the casein micelle that reduce the threat that lactation poses to the lifetime reproductive success of the mother: the control of ectopic calcification and the prevention of amyloidosis in the mammary gland. A third function is the need to form an easily digested gel in the stomach of the neonate. These three functions explain much of what we know about caseins and the structure of the casein micelle.

2. Origins and structures of casein genes

To quote the evolutionary biologist Theodosius Dobzhansky, “Nothing in biology makes sense except in the light of evolution” (Dobzhansky, 1973), a statement that is as true for caseins and the casein micelle as it is for anything else in biology.

Calcium phosphate mineralised tissues appeared more than 500 million years (My) ago. Concomitantly, a group of SCPPs (secreted, calcium (phosphate)-binding phosphoproteins) evolved (Kawasaki & Weiss, 2003). Members of this group are involved in every aspect of biomineralisation. The SCPP genes can be divided into two groups depending on whether or not they contain long exons encoding sequences that are rich in Pro and Gln residues (P,Q-rich sequences). The P,Q-rich sequences are sticky and encourage protein–protein interactions with low sequence specificity, sometimes called promiscuous interactions (Hsu et al., 2013; Kay, Williamson, & Sudol, 2000). Promiscuous interactions that involve only main-chain interactions are largely independent of sequence and are recognised to be important in, for example, the action of molecular chaperones and in enabling individual domains in a cell signalling network to respond to a variety of protein ligands (Macias, Wiesner, & Sudol, 2002).

Caseins all contain sticky, P,Q-rich, sequences, which is one reason why they readily associate with themselves and each other. Casein genes are all descendants of the gene ODAM “odontogenic ameloblast-associated protein” (Kawasaki, Lafont, & Sire, 2011), highly expressed in dental tissue, through two separate lineages. The κ-casein genes are derived from ODAM via FDCSP “follicular dendritic cell secreted peptide” that is active in the immune system and is highly expressed in adenoidal tissues. All other caseins, the so-called calcium-sensitive caseins, are derived from ODAM via the bone-associated protein SCPPPQ1 “SCPP P,Q-rich 1”. Calcium-sensitive caseins and peptides containing a phosphate centre can sequester amorphous calcium phosphate to form thermodynamically stable complexes (Little & Holt, 2004).

The other group of SCPPs, also known as SIBLINGS (small integrin-binding ligand, N-linked glycoproteins) (Fisher, Torchia, Fohr, Young, & Fedarko, 2001), are more acidic and do not contain sticky sequences. Like caseins they contain phosphate centre-type sequences and some have longer, highly phosphorylated, sequences. Among the earliest SCPPs to evolve was osteopontin (Kawasaki & Weiss, 2003). It is found in almost all species, tissues and biofluids and has at least 6 distinct functions (Mazzali et al., 2002; Scatena, Liaw, & Giachelli, 2007). It is the most abundant non-collagenous protein in the extracellular matrix of bone and is thought, as the name implies, to form a bridge between the mineral and osteocytes. A naturally occurring mixture of osteopontin phosphopeptides will form a type of calcium phosphate nanocluster in which the core of amorphous calcium phosphate is more basic, four times larger and more highly hydrated than the core of calcium phosphate nanoclusters sequestered by casein phosphopeptides (Holt, Sorensen, & Clegg, 2009). Chameleon SCPPPQ1 contains a casein phosphate-centre-type sequence, SASSSEE (http://www.uniprot.org/uniprot/E0YCE6) but in the rat (http://www.uniprot.org/uniprot/D6QY17) and mouse (http://www.uniprot.org/uniprot/B9UIU9) sequences it is modified to SGGSSSEQ (Kawasaki, 2009; Moffatt, Smith, Sooknanan, St-Arnaud, & Nanci, 2006). Phosphate centre-type sequences have been identified in three other SCPPs and in a number of other secreted phosphoproteins (Holt et al., 2009). The ability of caseins to sequester amorphous calcium phosphate is therefore shared with, and derived from, other SCPPs.

The first casein evolved more than 300 My ago in some stem amniote before the great divergence into synapsids, the mammalian lineage, and sauropsids, leading to birds, dinosaurs, turtles and crocodiles. Lactation probably originated in mammal-like reptiles such as cynodonts at least 50 My later (Lefèvre, Sharp, & Nicholas, 2010; Lemay et al., 2009; Oftedal, 2012). We have argued that the current biological function of caseins in milk is an adaptation of an antecedent function of caseins in the control of some aspect of biomineralisation (Holt & Carver, 2012). The four recognised casein orthologues were all established before true mammals evolved (Lefèvre et al., 2010; Rijnkels, Kooiman, de Boer, & Pieper, 1997). In the eutherian lineage a second type of α_{S2}-casein is found in some species and in the monotremes a second type of β-casein has evolved (Lefevre, Sharp, & Nicholas, 2009) so that there are six known casein gene products. All milks contain casein micelles. Notwithstanding this, there is considerable variation in the composition of caseins among extant mammals (Martin, Cebo, & Miranda, 2013). Only κ-casein is found in all species and individuals, along with at least two, and sometimes as many as four, other casein gene products. In eutherian species these other gene products are orthologues of the bovine α_{S1}-, α_{S2}- and β-caseins. Because the bovine orthologues can be precipitated by the addition of calcium ions at physiological pH, they are known as calcium-