ELSEVIER

The cryptome: a subset of the proteome, comprising cryptic peptides with distinct bioactivities

Dominic J. Autelitano[1], Antonio Rajic[1], A. Ian Smith[2], Michael C. Berndt[2], Leodevico L. Ilag[1] and Mathew Vadas[3]

[1] Cryptome Pharmaceuticals, PO Box 6492, St. Kilda Rd Central, Melbourne, Vic 8008, Australia
[2] Department of Biochemistry and Molecular Biology, Monash University, Clayton, Vic 3008, Australia
[3] Division of Human Immunology, Institute of Medical and Veterinary Science, PO Box 14, Rundle Mall, Adelaide SA 5000, Australia

There is increasing evidence that proteolytic cleavage gives rise to 'hidden' peptides with bioactivities that are often unpredicted and totally distinct to the parent protein. So far, the liberation of these cryptic peptides, or crypteins, has been shown to be prevalent in proteins associated with endocrine signalling, the extracellular matrix, the complement cascade and milk. A broad spectrum of proteases has been implicated in the generation of natural crypteins that appear to play a role in modulating diverse biological processes, such as angiogenesis, immune function and cell growth. The proteolytic liberation of crypteins with novel activities represents an important mechanism for increasing diversity of protein function and potentially offers new opportunities for protein-based therapeutics.

Before the exhaustive task of sequencing the entire human genome, it was expected that the complexity of an organism would be reflected by the number of genes encoding proteins or peptides. Various analyses predicted that the human genome would consist of up to 120,000 distinct genes [1]. However, subsequent sequence analysis demonstrated that the human genome contains <30,000 genes [2,3]. The genomes of the fruit fly *Drosophila melanogaster* and the mustard plant *Arabidopsis thaliana* contain ~13,600 and ~25,500 genes, respectively [4,5], indicating that the extent of biological intricacy apparent in higher organisms is not directly proportional to gene number, suggesting that additional factors must be considered.

Unlike the genome, the proteome is dynamic and reflects changes in cellular or tissue response to different conditions, including specific disease processes. It has been estimated that the human proteome could consist of as many as 1,000,000 distinct protein and peptide entities that are derived from the expression of <30,000 or so genes. Interestingly, there is evidence that the number of protein forms derived per gene increases as the complexity of the organism increases [6]. Extensive modulation of expression and function of proteins can occur at multiple levels ranging from transcription to post-translation. Post-transcriptional modifications, such as alternative splicing, use of alternative polyadenylation sites and RNA editing, can dramatically alter the diversity of protein products derived from a single gene. Post-translational modifications, including phosphorylation, glycosylation, acetylation, methylation and lipid attachment, can occur in various combinations to add further diversity to translated proteins and peptides [7].

However, potentially the most extensive range of post-translational modifications are generated by diverse groups of proteolytic enzymes that are collectively responsible for processing and/or cleavage of various classes of proteins into new fragments that often have unique biological activities. It is becoming increasingly obvious that proteases account for a significant proportion of the total genome, with current analysis suggesting that 2% of all human genes are proteases or protease inhibitors [8]. This provides an opportunity for generating a vast array of proteolytically cleaved peptides, many of which are as yet uncharacterized, thus adding considerable amplification and complexity to the range of bioactive peptides produced from a particular genome. There are many examples illustrating that one or more proteolytically cleaved peptide fragment(s) can be derived from a single protein precursor, and these peptides often show biological activities that

Corresponding author: Autelitano, D.J. (dautelitano@cryptomepharmaceuticals.com)

1359-6446/06/$ - see front matter © 2005 Elsevier Ltd. All rights reserved. doi:10.1016/j.drudis.2006.02.003

are distinct from the parent molecule. In many cases, these divergent biological activities are cryptic and cannot be predicted from either the amino acid sequence or the activity of the parent protein. The greatest number of cryptic activities discovered so far is derived via cleavage of extracellular matrix and milk proteins. It is probable that the cryptic matrix-derived peptides play an important local role as modulators of angiogenesis and tissue remodelling, whereas bioactivities encrypted in milk proteins could have diverse immunomodulatory and antimicrobial roles in the neonate and the adult.

It is becoming more apparent that a wide range of proteins contain concealed functional units that can be liberated under certain conditions to generate novel bioactivities. We propose that this cryptic subset of peptides, residing within the proteome, is termed the cryptome and that it potentially represents a vast array of cryptic peptides, or crypteins, with varying bioactivities that can be liberated from the parent protein via proteolytic cleavage. This review will focus on several examples of such cryptic peptide activities and will highlight the potential therapeutic opportunities for this diverse group of molecules.

Diverse biological activities from a single translation product; proopiomelanocortin as a model prohormone

The generation of biologically active peptides and polypeptides generated via cleavage of precursor proteins is a common phenomenon that is widely used throughout nature. One of the most striking and best studied examples is that of the prohormone precursor proopiomelanocortin (POMC), which gives rise to an array of diverse bioactive peptides (Figure 1). Characterization of the structural relationships between adrenocorticotropic hormone (ACTH) and β-lipotrophin (β-LPH) showed that these distinct biologically active peptides were derived from a common precursor protein, POMC [9]. The full extent of the complexity of this single translation product became apparent following the cloning and sequence analysis of POMC and it is now well-established that this prohormone not only gives rise to ACTH and β-LPH, but also to additional bioactive peptides including β-endorphin, α-, β- and γ-melanocyte-stimulating hormone (MSH), corticotrophin-like intermediate lobe peptide (CLIP), N-terminal peptide of proopiomelanocortin (N-POMC) and joining peptide (JP) [10]. All of these

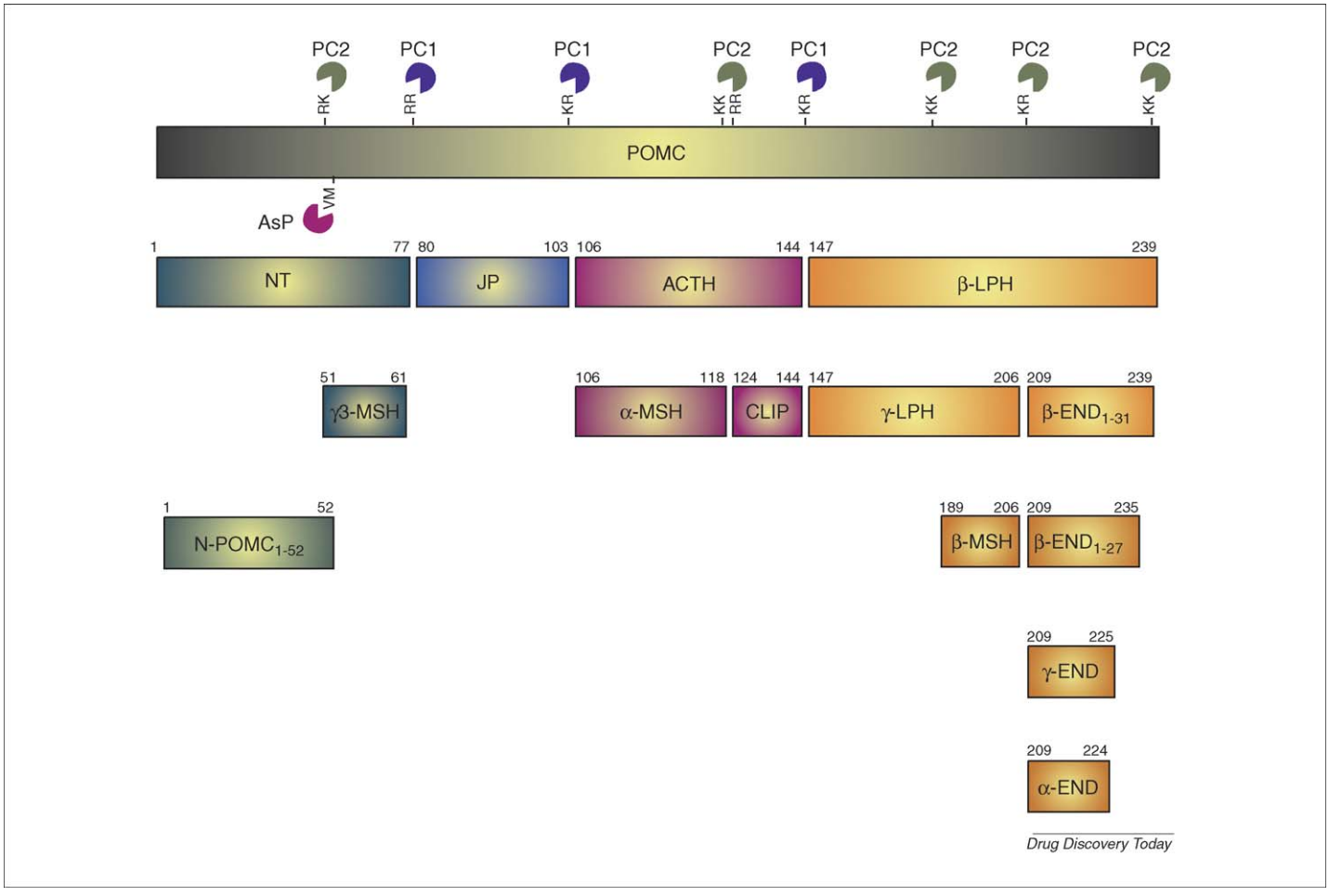

FIGURE 1

Schematic representation of the processing of the multifunctional prohormone pro-opiomelanocortin. Differential proteolytic processing of the initial translation product occurs as a result of the tissue-specific expression of the prohormone convertases PC1 and PC2, as well as other proteases such as AsP. Proteolytic cleavage of POMC gives rise to >12 distinct peptides with diverse functions, thus generating substantial amplification of the biological potential of a single protein. The known cleavage sites are indicated by the paired amino acids and the numbers refer to the amino acid positions of the liberated peptides. Abbreviations: POMC, proopiomelanocortin; NT, N-terminal peptide; JP, joining peptide; ACTH, adrenocorticotropic hormone; LPH, lipotropin; MSH, melanocyte-stimulating hormone; CLIP, corticotropin-like intermediate peptide; END, endorphin; PC1, prohormone convertase 1; PC2, prohormone convertase 2; AsP, adrenal secretory protease; R, arginine; K, lysine; V, valine; M, methionine.

peptides are involved in a plethora of biological responses with physiological or pathophysiological relevance [11].

Because expression of the POMC gene is widespread and occurs in the brain and pituitary, as well as many peripheral tissues, additional mechanisms direct the selected expression of individual bioactive POMC-derived peptides in a cell- or tissue-specific manner [10,11]. Cell-specific release of different POMC-derived peptides was shown to be conferred by the two prohormone convertases PC1 and PC2, which are subtilisin-like endoproteases with distinct substrate specificities [12]. PC1 and PC2, acting either alone or sequentially, can generate all of the physiologically relevant POMC-derived peptides in the brain and pituitary [12]. Furthermore, the cell-specific expression of these prohormone convertases results in the PC1-mediated production of ACTH and β-LPH in the anterior pituitary corticotroph, whereas co-expression of PC1 and PC2 in pituitary melanotrophs and hypothalamic neurons leads to further processing and the release of α-MSH, CLIP and β-endorphin [11,13]. Further biodiversity is added to the β-endorphin-related peptides via carboxypeptidase shortening and N-acetylation, resulting in altered biological functions, particularly with respect to opiate activity [10]. Additional enzymes, such as the adrenal secretory protease (AsP) described by Bicknell *et al.* [14], cleave POMC in a tissue-restricted fashion to generate the adrenal mitogenic peptide N-POMC$_{1–52}$.

The differential processing of a single translation product to generate multiple peptide bioactivities is clearly not limited to POMC and has been unequivocally demonstrated with other prohormones, including the opioid precursors proenkephalin and prodynorphin [15], prosomatostatin [16] and prothyrotropin-releasing hormone [17], suggesting that this represents a common and frugal approach to increasing the biodiversity of the proteome.

PC1 and PC2 are part of a larger family of prohormone convertases (also comprising furin, PC4, PC5, PACE4 and PC7) that act either alone or in various combinations to liberate many bioactive peptides from various prohormones in a tissue-specific manner by cleaving at single or paired basic residues [13]. Collectively, this group of proteases is responsible for functional processing of not only prohormones, but also precursor forms of proteolytic enzymes, growth factors, signalling molecules, transcription factors, adhesion molecules and cell surface proteins of human pathogens [13]. Thus, it is obvious that even this small group of proteases with limited selectivities can create considerable amplification of biological signals by acting either individually or sequentially on a limited number of primary translation products, generating new peptides with diverse biological capabilities.

Extracellular-matrix-derived crypteins

Proteins contributing to the extracellular matrix (ECM) provide not only structural support, but also play a major role in coordinating various cell signalling events. There is now overwhelming evidence that remodelling of the vascular basement membrane ECM serves not only to break down the physical matrix barrier in preparation for endothelial cell migration, but also to expose a multitude of otherwise obscured, cryptic proteins that play a role in modulating the angiogenic response to tissue injury and wound healing, as well as the pathological response to tumour cell growth [18–20].

In addition to their structural properties, the collagen α chains have been unequivocally shown to contain several cryptic polypeptides that can be released from the ECM to act as modulators of angiogenesis and tumour cell growth (Figure 2). A 20 kDa C-terminal fragment (named endostatin), derived from the non-collagenous domain of collagen XVIII, was shown to inhibit angiogenesis and the *in vivo* growth of primary and metastatic tumours, demonstrating that this cleaved peptide fragment had biological functions distinct from the parent protein [21]. A 22 kDa anti-angiogenic peptide, restin, was derived from the noncollagenous domain of human collagen XV and was shown to have different anti-angiogenic actions to endostatin [22,23].

Type IV collagen is one of the major structural constituents of basement membrane ECM and is expressed in six distinct α chains (α_1–α_6) that are assembled into triple helical bundles. The C-terminal noncollagenous domains of several of the type IV collagen α chains have also been shown to possess potent anti-angiogenic and antitumour properties. Arresten represents a 26 kDa noncollagenous1 (NC1) domain fragment of the collagen type IV α_1 chain [24], canstatin is a fragment of the NC1 domain of the collagen type IV α_2 chain [25] and tumstatin is a fragment of the NC1 domain of the collagen type IV α_3 chain [26,27]. All of these peptides, derived from the C-terminal domains of collagen type IV α chains, have demonstrated direct anti-angiogenic effects that are mediated by distinct interactions with integrins [27]. Furthermore, the noncollagenous domain of the α_3 chain of type IV collagen (tumstatin) contains a peptide domain (amino acids 54–132) that is responsible for anti-angiogenic activity [26], as well as a distinct peptide (amino acids 185–203) that can directly inhibit tumour cell proliferation [28]. Comprehensive reviews on the generation and functions of various collagen-derived crypteins as biological regulators of angiogenesis have been published previously [29–31].

Several proteolytic enzymes, including various matrix metalloproteinases (MMPs), serine proteases and cathepsins, are capable of generating cryptic bioactive peptides from ECM components such as the collagens, indicating that these peptides can be generated *in vivo* in a regulated fashion and might then act locally as modulators of angiogenesis [32–34]. Collectively, these data suggest that these biologically active crypteins are generated by proteolytic cleavage and that different enzymes might be capable of generating similar and/or overlapping peptide fragments in a cell- or tissue-specific manner.

The generation of biologically active crypteins from the ECM is not limited to the family of collagens [19,31,35]. Endorepellin is a C-terminal fragment of perlecan, a basement membrane proteoglycan that also functions as a potent angiogenesis inhibitor [36]. Release of the laminin-5 γ_2-chain-derived domain III peptide by membrane type1 (MT1)-MMP- and MMP-2-mediated proteolysis generates a novel peptide with the ability to activate the epidermal growth factor (EGF) receptor and to induce epithelial cell migration [19,37,38], an activity that is not apparent in the parent molecule. An unrelated cryptic peptide derived from the α_5 chain of laminin-10 was shown to induce inflammatory cell chemotaxis *in vitro* and *in vivo*, as well as inducing metalloproteinase activity and cytokine activation [39].

Other endogenous crypteins that regulate angiogenesis

The generation of cryptic fragments that modulate angiogenesis appears to be a common phenomenon that is not limited to

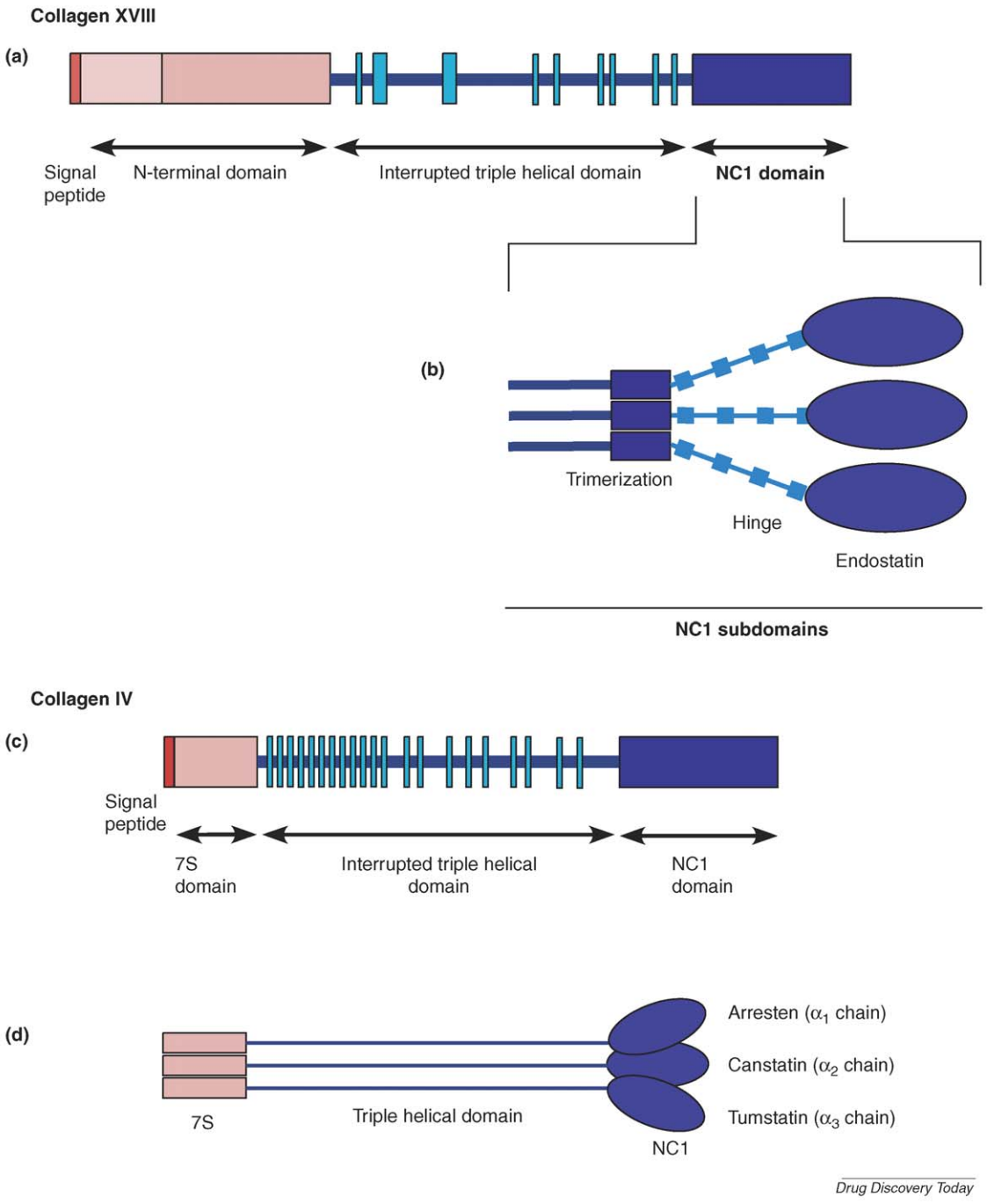

FIGURE 2

Structural domains and cryptic peptides of mammalian collagen XVIII and collagen IV. (a) A single chain of collagen XVIII comprises an N-terminal domain, a central domain made up of multiple triple helical and noncollagenous (NC) regions and a C-terminal NC1 domain. **(b)** The NC1 domain of collagen XVIII is composed of subdomains, including a region responsible for trimerization, a hinge domain that is the target of various proteases and the anti-angiogenic component endostatin that can be liberated after cleavage. **(c)** The linear structure of a single collagen IV α_1 chain is depicted, showing the typical cysteine-rich N-terminal 7S domain, the central triple-helical domain interrupted by small NC regions and a globular C-terminal NC1 domain. **(d)** Triple helical bundles of collagen IV are assembled via an interaction between the NC1 domains and covalent association of the 7S domains. Several proteases can act to liberate cryptic anti-angiogenic fragments, such as arresten, canstatin and tumstatin, from NC1 domains of the various α chains.

ECM-derived protein substrates. Examples include angiostatin, a 38 kDa fragment of plasminogen that blocks neovascularization and growth of metastatic tumours [40], and the 50 kDa plasmin-cleavage product of fibrinogen, which inhibits endothelial cell migration and tubule formation [41]. High-molecular-weight kininogen (HK) is processed by plasma kallikrein to generate bradykinin and HKa that have opposing biological effects on angiogenesis [42]. Further proteolytic processing of bradykinin

by angiotensin-converting enzyme (ACE) abrogates its pro-angiogenic effects by generating thrombostatin, a peptide that inhibits α-thrombin-induced platelet activation [43]. Multiple members of the prolactin–growth hormone family contain 16 kDa N-terminal domain peptides that, after proteolytic processing, inhibit *in vitro* endothelial cell growth, tube formation and *in vivo* angiogenesis [44,45].

There is also evidence that proteins associated with discrete intracellular locations can be proteolytically processed to generate cryptic fragments that are then secreted. Vasostatin, a potent endothelial cell inhibitor purified from the supernatant of Epstein–Barr virus (EBV)-immortalized cells, was identified as the N-terminal fragment of the Ca^{2+}-binding protein calreticulin, which is localized within the endoplasmic reticulum [46].

Human tyrosyl-tRNA synthetase (TyrRS) is secreted from cells undergoing apoptosis, and is cleaved by leukocyte elastase into two distinct fragments [47]: a 40 kDa N-terminal fragment of TyrRS stimulates endothelial cell chemotaxis and angiogenesis, as well as acting as an IL-8-like polymorphonuclear leukocyte chemokine [47,48]; and a 24 kDa C-terminal fragment with distinct chemotactic and activating effects on leukocytes [47]. By contrast, a fragment derived from tryptophanyl-tRNA synthetase (TrpRS), a close homologue of TyrRS, was shown to be angiostatic *in vitro* and *in vivo* [49,50], as a result of binding to vascular endothelial (VE)-cadherin at intercellular junctions of endothelial cells [51].

Milk crypteins

Milk contains a multitude of bioactive proteins that, in addition to their well-characterized functional roles, give rise to cryptic peptide activities that become apparent only after proteolytic digestion (reviewed in [52]). In particular, the caseins, β-lactoglobulin and α-lactalbumin have been shown to give rise to cryptic peptides with diverse bioactivities, such as immunomodulatory, antihypertensive, antimicrobial and mineral-binding activities [52]. These cryptic bioactivities are generated *in vivo* by the action of either digestive proteases or proteases associated with microbial fermentation of milk proteins, suggesting that these peptides could indeed play a physiological role.

For example, peptidic fractions generated during *Lactobacillus* fermentation of milk have been shown to possess immunomodulatory bioactivities [53,54], as well as to promote bone formation in osteoblast cultures [55]. Peptides derived by proteolytic fragmentation of α- and β-caseins, α-lactalbumin and β-lactoglobulin have the ability to interact with various opiate receptor subtypes, both as agonists and antagonists [56]. Some casein-derived casomorphins inhibit growth of human breast cancer cells [57] and promote neurite outgrowth in mouse neuroblastoma cells [58].

Pepsin digestion of bovine lactoferrin generates a peptide (Lfn-p) that induces apoptotic death of human oral squamous carcinoma cells *in vitro* [59], whereas tryptic digestion of the caseins yields several phosphopeptides that can complex with calcium phosphate. These caseinophosphopeptides act as a reservoir for biologically available calcium phosphate and have the ability to buffer tooth plaque pH and promote tooth enamel remineralization in animal studies and human trials [60,61]. Moreover, several studies have conclusively demonstrated that various caseinophosphopeptides are generated *in vivo* following ingestion of milk, yogurt or casein, and that some of these peptides appear to be absorbed and can be detected in plasma [62,63].

A further well-described cryptic function of proteolytic peptide fragments derived from milk proteins is the hypotensive action that appears to be caused by inhibiting ACE activity (reviewed in [64]). Casokinins and lactokinins (ACE-inhibitory peptides) are generated via the digestion of milk-derived caseins and β-lactoglobulin [65,66].

Classes of crypteins

The examples previously cited demonstrate clearly that, within the proteome, a subset of cryptic peptides exist whose biological activity only becomes apparent after proteolytic cleavage from the parent molecule. We refer to these peptides and/or polypeptides as crypteins and propose that several classes of crypteins exist, either as naturally processed fragments or, alternatively, as the result of artificial proteolytic cleavage (Box 1). An overview of the characteristics and derivation of several well-defined crypteins is given in Table 1. We propose that type 1 crypteins are produced when a parent protein is naturally cleaved by one or more proteases to generate a peptide or group of peptides with novel bioactivities that are distinct from the source protein. Many of the cryptic peptides described in this review fall into this category, with the anti-angiogenic peptides derived from collagen α chains, calreticulin or plasminogen being obvious examples. Interestingly, despite their obvious overlap in terms of biological effect, collagens, calreticulin and plasminogen have little in common to suggest that they might harbour inherent, but cryptic, anti-angiogenic activities. Moreover, these crypteins do not necessarily show significant amino acid sequence homology nor highly related structural features, which could explain their biological activity. Thus, this cryptic subset of bioactivities that resides within the proteome is difficult to predict by modern proteomic approaches that rely on sequence and/or structural comparisons.

We define type 2 crypteins as naturally liberated fragments of a parent protein that either retain the original or a related bioactivity of the parent molecule. This type of cryptein could be a peptide that encompasses the bioactive motif only. An example is the insulin-like growth factor-1 (IGF-1)-derived N-terminal tripeptide (glycine-proline-glutamate) that appears to be naturally generated in plasma and brain tissue [67,68]. Although the tripeptide derivative has a similar neuroprotective activity to the parent protein, its ability to cross the blood–brain barrier is enhanced, thereby providing specific advantages for uptake and transport into central nervous system (CNS) tissue [67,68]. Thus, the proteolytic generation of this tripeptide could be part of a

BOX 1

Classes of crypteins

Type 1: natural proteolytic cleavage of parent protein yields product with novel, unrelated bioactivity.

Type 2: natural proteolytic cleavage of parent protein yields product with similar bioactivity to parent molecule but might have modified features.

Type 3: *in vitro* generated protein–peptide fragments with novel bioactivities. Identical or similar fragments might not necessarily be generated naturally.

TABLE 1

Examples of crypteins

Cryptein	Cryptein derived from	Bioactivity of cryptein	Cryptein generated by	Cryptein type
Angiostatin [40]	38 kDa plasminogen fragment	Anti-angiogenic Inhibits tumour cell growth Inhibits tumour cell	Elastase Matrix metalloproteinase (MMP)-2,-3,-7 and -9 (reviewed in [81])	1
Endostatin [21]	20 kDa C-terminal fragment of collagen XVIII	Anti-angiogenic Inhibits tumour cell growth	Elastase Cathepsin-L and -B MMP-7 (reviewed in [31])	1
Tumstatin [26]	28 kDa fragment of the NC1 domain of the collagen type IV α_3 chain	Anti-angiogenic and antitumour activity associated with distinct domains	Predominantly MMP-9, also MMP-2,-3 and -13 (reviewed in [31])	1
Endorepellin [36]	85 kDa C-terminal fragment of perlecan	Anti-angiogenic	Possible perlecan cleavage by plasmin, thrombin, MMP-1 and -2 [82]	1
Vasostatin [46]	N-terminal fragment of calreticulin	Anti-angiogenic	Unknown	1
N-terminal prolactin [44]	16 kDa N-terminal fragment of prolactin	Anti-angiogenic	Unknown	1
N-terminal tyrosyl-tRNA synthetase [47]	40 kDa fragment of human tyrosyl-tRNA synthetase	Angiogenic [48]; polymorphonuclear leukocyte chemotaxis [47]	Elastase [47]	1
DIII peptide of laminin-5 [19]	Laminin-5 γ_2 chain	Epidermal growth factor receptor activation Epithelial cell migration	Membrane type 1 (MT1)-MMP, MMP-2 (reviewed in [19])	1
Laminin-10-derived chemotactic peptide [39]	Laminin-10 α_5 chain	Inflammatory cell chemotaxis and activation	Unknown	1
Thrombostatin [43]	Bradykinin	Inhibitor of α-thrombin-induced platelet activation	Angiotensin-converting enzyme [43]	1
Casocidin-I [83]	39-amino-acid peptide derived from α_{s2}-casein	Antibacterial agent	Likely to be derived by digestive or bacterial proteases [83]	1
β-lactoglobulin peptides [54]	Several peptides from milk-derived β-lactoglobulin	Immunomodulatory	*Lactobacillus*-derived proteases [54]	1
N-terminal platelet factor-4 [69]	Platelet factor-4	Endothelial cell-growth-modulator with enhanced inhibitory action	Unknown	2
Glypromate (Gly-Pro-Glu) [68]	Tripeptide derived from N-terminus of insulin-like growth factor-1	Neuroprotectant peptide that can cross blood–brain barrier	Unknown	2
AOD9604 [72]	C-terminus of human growth hormone	Modulates lipid metabolism	Unknown	3
Caseinophosphopeptides [60]	Various peptides derived from caseins	Complex with calcium phosphate – promotes tooth-enamel remineralization	Trypsin (*in vitro*) Digestive or bacterial proteases (*in vivo*) [60,62]	3

physiological or pathophysiological mechanism that allows access of this neuroprotectant molecule to the CNS.

A further example of this type of cryptein (in which additional properties are gained following proteolytic cleavage) is illustrated by the N-terminal-processed fragment of platelet factor-4 (PF-4) [69], which is 50 times more potent than intact PF-4 in inhibiting endothelial cell proliferation. Thus, despite the apparent redundancy in generating small crypteins with similar properties to the parent protein, it is likely that additional physiologically relevant properties could be gained, including altered affinity, stability and tissue availability.

In many instances, the precise proteolytic enzymes involved in generating type 2 crypteins have not been fully elucidated; however, it is conceivable that similar fragments with indistinguishable biological effects can result from cleavage with more than one protease. This suggests that in some cases precise cleavage might not be an absolute requirement and liberation of a particular motif is sufficient to unmask otherwise unrecognized cryptic functions. It is possible to exploit this by artificially generating a range of distinct and overlapping peptides *in vitro*, using available proteases to generate cryptic fragments (type 3 cryptein) that might not necessarily represent the precise biological products generated naturally. Examples of this type of cryptein are the caseinophosphopeptides that can be generated via *in vitro* digestion of casein with trypsin [60] and the apoptotic cell-death-inducing peptide (Lfn-p) that is cleaved from lactoferrin by pepsin digestion [59]. Although these peptides have been generated *in vitro*, it is likely that biologically similar peptides are released following *in vivo* digestion of milk-related products [63]. A systematic approach to the *in vitro* generation of cryptic peptides (derived from a wide range of proteins and/or biological extracts) has been established and has the potential to uncover previously unrecognized bioactive peptides with potential therapeutic applications (PCT: WO 2004/008148).

Crypteins as therapeutics

Biopharmaceutical drugs, comprising synthetic and recombinant proteins, monoclonal antibodies and nucleic-acid-based compounds, have received considerable attention over recent years with an increasing number of these products entering the market place. During the period 2000–2003, the biopharmaceutical market underwent steady growth, with over a quarter of new drug approvals being biopharmaceuticals [70]. Regulatory authorities in North America and Europe approved a total of 64 protein-based biopharmaceuticals (comprising 30 new products) for human use during this three year period; another 500 biopharmaceuticals were in stages of clinical development [70]. The benefits of protein-based drugs include potentially low toxicity and high specificity that might allow the candidate to move through clinical development faster than small-molecule drugs [71]. In addition, protein and/or peptide drugs can potentially cover broad areas of therapeutic applications, such as diabetes, inflammation, cardiovascular disease and cancer, making them attractive lead compounds.

As a class of protein-based molecules, crypteins offer a novel perspective on developing new therapeutic agents with a broad range of bioactivities. As more bioactive crypteins are unmasked, the range of functional proteins that might have useful interactions with therapeutic targets should also increase, thereby potentially providing novel opportunities for drug development. In some instances, smaller cryptic fragments might provide better drug candidates than their parent proteins because of enhanced specificity, delivery and ease of production. Several of the cryptic proteins described in this review have been exploited as therapeutic agents, be it as unmodified synthetic peptides, recombinant proteins or after some form of engineering that tailors specific properties to suit the therapeutic need.

A natural proteolytically cleaved fragment of IGF-1 (glycine-proline-glutumate) with potent neuroprotectant properties [68] has recently been fast-tracked by the FDA to Phase III trials in coronary artery bypass graft surgery, highlighting the potential of such molecules in areas where new therapeutic approaches are desperately needed. Similarly, a peptide derived from α-MSH (CZEN-002) has recently shown positive effects in early trials as an antifungal agent with minimal adverse reactions, potentially offering new options to azole-based antifungal drugs (www.zengen.com). A small peptide cryptein corresponding to the C-terminus of human growth hormone was shown to mediate the antilipogenic effects of growth hormone [72] and is currently undergoing Phase II clinical trials as an antiobesity drug (www.metabolic.com.au). Several clinical trials, including the recent study reported by Mizuno *et al.* [73], have demonstrated significant, dose-dependent lowering of blood pressure after oral administration of casein-derived peptide fragments that are thought to act as ACE inhibitors.

Cryptic peptides with novel biological activities are also beginning to find beneficial therapeutic uses as food additives. In particular, the trypsin-derived caseinophosphopeptides have demonstrated beneficial effects on the remineralization of tooth enamel when simply added to chewing gum and even to sports drinks [74,75].

In addition to small peptide crypteins as drugs, larger crypteins have also been explored as potential therapeutic agents. Angiostatin and endostatin have been tested in Phase I and II trials as potential antitumour agents. Phase I studies with endostatin, initiated by EntreMed, demonstrated that it was well-tolerated with little or no adverse effects, and some evidence of clinical benefit was suggested at the doses administered [76]. In the future, potential problems associated with the manufacture and administration of the original 20 kDa endostatin [77] could be aided by the recent demonstration that all of the anti-angiogenic and antitumour activities are localized to a 27-amino-acid N-terminal peptide [78], potentially providing a simpler and a more cost-effective strategy for manufacture. Furthermore, it might be feasible to enhance the therapeutic potency of endostatin using a recently described single amino acid modified version that is more effective *in vitro* and *in vivo* [79,80].

The potential benefit of angiostatin in combination with other anticancer drugs is also being assessed (ClinicalTrials.gov Identifier: NCT00049790) and this could provide additional advantages over current strategies. In addition, an anti-angiogenic recombinant protein derived from the NC1 domain of the α_2 chain of human collagen type IV (Angiocol™) has received FDA approval to commence trials in cancer patients (www.biostratum.com).

Conclusions

The generation of new peptides and proteins via proteolytic cleavage is a fundamental and common process used by organisms

to modify primary translation products as a means of either activating or deactivating a specific activity, or to liberate an otherwise biologically inert peptide. It is becoming apparent that many proteins carry multiple hidden functionalities that can be unmasked following proteolytic cleavage. The existence of these cryptic bioactivities, buried within a protein sequence, appears to be a common theme and could reflect an evolutionary mechanism designed to expand the functionality of proteins. Increased recognition of new, cryptic bioactivities will no doubt provide additional opportunities for understanding and modulating various biological processes, and in some cases might present novel ligands that could provide functional therapeutic tools. The ECM-derived angiogenesis inhibitors and milk-protein-derived ACE inhibitors provide examples of cryptic peptides that have provided novel ligands that are currently being evaluated as potential therapeutics. Thus, newly discovered cryptic peptide ligands might provide additional strategies for the therapeutic modulation of a variety of target indications and, because they are derived from endogenous, natural protein sequences, could offer relatively safe therapeutic alternatives. Like other candidates in the growing area of protein therapeutics, these compounds will need to tackle the obvious challenges in drug development, particularly those relating to delivery and manufacturing. Despite these challenges, several of the crypteins described in this review, particularly the smaller peptides (such as glypromate, the α-MSH-derived antimicrobial agents and the casein-derived tripeptide ACE inhibitors), are currently progressing through clinical trials and appear to offer realistic therapeutic alternatives.

Acknowledgements

The authors wish to thank Leanne Neale for critical reading of the manuscript.

References

1 Liang, F. *et al.* (2000) Gene index analysis of the human genome estimates approximately 120,000 genes. *Nat. Genet.* 25, 239–240
2 Lander, E.S. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
3 Venter, J.C. *et al.* (2001) The sequence of the human genome. *Science* 291, 1304–1351
4 Adams, M.D. *et al.* (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195
5 Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815
6 Wilkins, M.R. *et al.* (1996) Current challenges and future applications for protein maps and post-translational vector maps in proteome projects. *Electrophoresis* 17, 830–838
7 Subramanian, G. *et al.* (2001) Implications of the human genome for understanding human biology and medicine. *JAMA* 286, 2296–2307
8 Puente, X.S. *et al.* (2005) A genomic view of the complexity of mammalian proteolytic systems. *Biochem. Soc. Trans.* 33, 331–334
9 Mains, R.E. *et al.* (1977) Common precursor to corticotropins and endorphins. *Proc. Natl. Acad. Sci. USA* 74, 3014–3018
10 Smith, A.I. and Funder, J.W. (1988) Proopiomelanocortin processing in the pituitary, central nervous system, and peripheral tissues. *Endocr. Rev.* 9, 159–179
11 Raffin-Sanson, M.L. *et al.* (2003) Proopiomelanocortin, a polypeptide precursor with multiple functions: from physiology to pathological conditions. *Eur. J. Endocrinol.* 149, 79–90
12 Benjannet, S. *et al.* (1991) PC1 and PC2 are proprotein convertases capable of cleaving proopiomelanocortin at distinct pairs of basic residues. *Proc. Natl. Acad. Sci. USA* 88, 3564–3568
13 Seidah, N.G. and Chretien, M. (1999) Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. *Brain Res.* 848, 45–62
14 Bicknell, A.B. *et al.* (2001) Characterization of a serine protease that cleaves pro-gamma-melanotropin at the adrenal to stimulate growth. *Cell* 105, 903–912
15 Zamir, N. *et al.* (1984) Differential processing of prodynorphin and proenkephalin in specific regions of the rat brain. *Proc. Natl. Acad. Sci. USA* 81, 6886–6889
16 Benoit, R. *et al.* (1990) Processing of prosomatostatin. *Metabolism* 39 (Suppl. 2), 22–25
17 Ladram, A. *et al.* (1994) Modulation of the biological activity of thyrotropin-releasing hormone by alternate processing of pro-TRH. *Biochimie* 76, 320–328
18 Davis, G.E. *et al.* (2000) Regulation of tissue injury responses by the exposure of matricryptic sites within extracellular matrix molecules. *Am. J. Pathol.* 156, 1489–1498
19 Schenk, S. and Quaranta, V. (2003) Tales from the crypt[ic] sites of the extracellular matrix. *Trends Cell Biol.* 13, 366–375
20 Sage, E.H. (1997) Pieces of eight: bioactive fragments of extracellular proteins as regulators of angiogenesis. *Trends Cell Biol.* 7, 182–186
21 O'Reilly, M.S. *et al.* (1997) Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. *Cell* 88, 277–285
22 Ramchandran, R. *et al.* (1999) Antiangiogenic activity of restin. NC10 domain of human collagen XV: comparison to endostatin. *Biochem. Biophys. Res. Commun.* 255, 735–739
23 Sasaki, T. *et al.* (2000) Endostatins derived from collagens XV and XVIII differ in structural and binding properties, tissue distribution and anti-angiogenic activity. *J. Mol. Biol.* 301, 1179–1190
24 Colorado, P.C. *et al.* (2000) Anti-angiogenic cues from vascular basement membrane collagen. *Cancer Res.* 60, 2520–2526
25 Kamphaus, G.D. *et al.* (2000) Canstatin, a novel matrix-derived inhibitor of angiogenesis and tumor growth. *J. Biol. Chem.* 275, 1209–1215
26 Maeshima, Y. *et al.* (2000) Distinct antitumor properties of a type IV collagen domain derived from basement membrane. *J. Biol. Chem.* 275, 21340–21348
27 Petitclerc, E. *et al.* (2000) New functions for non-collagenous domains of human collagen type IV. Novel integrin ligands inhibiting angiogenesis and tumor growth *in vivo*. *J. Biol. Chem.* 275, 8051–8061
28 Han, J. *et al.* (1997) A cell binding domain from the alpha3 chain of type IV collagen inhibits proliferation of melanoma cells. *J. Biol. Chem.* 272, 20395–20401
29 Marneros, A.G. and Olsen, B.R. (2001) The role of collagen-derived proteolytic fragments in angiogenesis. *Matrix Biol.* 20, 337–345
30 Ortega, N. and Werb, Z. (2002) New functional roles for non-collagenous domains of basement membrane collagens. *J. Cell Sci.* 115, 4201–4214
31 Bix, G. and Iozzo, R.V. (2005) Matrix revolutions: "tails" of basement-membrane components with angiostatic functions. *Trends Cell Biol.* 15, 52–60
32 Xu, J. *et al.* (2001) Proteolytic exposure of a cryptic site within collagen type IV is required for angiogenesis and tumor growth *in vivo*. *J. Cell Biol.* 154, 1069–1079
33 Felbor, U. *et al.* (2000) Secreted cathepsin L generates endostatin from collagen XVIII. *EMBO J.* 19, 1187–1194
34 Ferreras, M. *et al.* (2000) Generation and degradation of human endostatin proteins by various proteinases. *FEBS Lett.* 486, 247–251
35 Mott, J.D. and Werb, Z. (2004) Regulation of matrix biology by matrix metalloproteinases. *Curr. Opin. Cell Biol.* 16, 558–564
36 Mongiat, M. *et al.* (2003) Endorepellin, a novel inhibitor of angiogenesis derived from the C terminus of perlecan. *J. Biol. Chem.* 278, 4238–4249
37 Koshikawa, N. *et al.* (2000) Role of cell surface metalloprotease MT1-MMP in epithelial cell migration over laminin-5. *J. Cell Biol.* 148, 615–624
38 Giannelli, G. *et al.* (1997) Induction of cell migration by matrix metalloprotease-2 cleavage of laminin-5. *Science* 277, 225–228
39 Adair-Kirk, T.L. *et al.* (2003) A site on laminin alpha 5, AQARSAASKVKVSMKF, induces inflammatory cell production of matrix metalloproteinase-9 and chemotaxis. *J. Immunol.* 171, 398–406
40 O'Reilly, M.S. *et al.* (1994) Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* 79, 315–328
41 Bootle-Wilbraham, C.A. *et al.* (2000) Fibrinogen E-fragment inhibits the migration and tubule formation of human dermal microvascular endothelial cells *in vitro*. *Cancer Res.* 60, 4719–4724

42 Guo, Y.L. and Colman, R.W. (2005) Two faces of high-molecular-weight kininogen (HK) in angiogenesis: bradykinin turns it on and cleaved HK (HKa) turns it off. *J. Thromb. Haemost.* 3, 670–676

43 Prieto, A.R. *et al.* (2002) Thrombostatin, a bradykinin metabolite, reduces platelet activation in a model of arterial wall injury. *Cardiovasc. Res.* 53, 984–992

44 Ferrara, N. *et al.* (1991) The 16K fragment of prolactin specifically inhibits basal or fibroblast growth factor stimulated growth of capillary endothelial cells. *Endocrinology* 129, 896–900

45 Struman, I. *et al.* (1999) Opposing actions of intact and N-terminal fragments of the human prolactin/growth hormone family members on angiogenesis: an efficient mechanism for the regulation of angiogenesis. *Proc. Natl. Acad. Sci. USA* 96, 1246–1251

46 Pike, S.E. *et al.* (1998) Vasostatin, a calreticulin fragment, inhibits angiogenesis and suppresses tumor growth. *J. Exp. Med.* 188, 2349–2356

47 Wakasugi, K. and Schimmel, P. (1999) Two distinct cytokines released from a human aminoacyl-tRNA synthetase. *Science* 284, 147–151

48 Wakasugi, K. *et al.* (2002) Induction of angiogenesis by a fragment of human tyrosyl-tRNA synthetase. *J. Biol. Chem.* 277, 20124–20126

49 Otani, A. *et al.* (2002) A fragment of human TrpRS as a potent antagonist of ocular angiogenesis. *Proc. Natl. Acad. Sci. USA* 99, 178–183

50 Wakasugi, K. *et al.* (2002) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. *Proc. Natl. Acad. Sci. USA* 99, 173–177

51 Tzima, E. *et al.* (2005) VE-cadherin links tRNA synthetase cytokine to anti-angiogenic function. *J. Biol. Chem.* 280, 2405–2408

52 Meisel, H. (2004) Multifunctional peptides encrypted in milk proteins. *Biofactors* 21, 55–61

53 Leblanc, J. *et al.* (2004) Induction of a humoral immune response following an *Escherichia coli* O157:H7 infection with an immunomodulatory peptidic fraction derived from Lactobacillus helveticus-fermented milk. *Clin. Diagn. Lab. Immunol.* 11, 1171–1181

54 Prioult, G. *et al.* (2004) Stimulation of interleukin-10 production by acidic beta-lactoglobulin-derived peptides hydrolyzed with Lactobacillus paracasei NCC2461 peptidases. *Clin. Diagn. Lab. Immunol.* 11, 266–271

55 Narva, M. *et al.* (2004) Effects of *Lactobacillus helveticus* fermented milk on bone cells *in vitro*. *Life Sci.* 75, 1727–1734

56 Teschemacher, H. *et al.* (1997) Milk protein-derived opioid receptor ligands. *Biopolymers* 43, 99–117

57 Kampa, M. *et al.* (1996) Identification of a novel opioid peptide (Tyr-Val-Pro-Phe-Pro) derived from human alpha S1 casein (alpha S1-casomorphin, and alpha S1-casomorphin amide). *Biochem. J.* 319, 903–908

58 Sakaguchi, M. *et al.* (2003) Neurite outgrowth-stimulating activities of beta-casomorphins in Neuro-2a mouse neuroblastoma cells. *Biosci. Biotechnol. Biochem.* 67, 2541–2547

59 Sakai, T. *et al.* (2005) Pepsin-digested bovine lactoferrin induces apoptotic cell death with JNK/SAPK activation in oral cancer cells. *J. Pharmacol. Sci.* 98, 41–48

60 Reynolds, E.C. (1998) Anticariogenic complexes of amorphous calcium phosphate stabilized by casein phosphopeptides: a review. *Spec. Care Dentist* 18, 8–16

61 Reynolds, E.C. *et al.* (1995) Anticariogenicity of calcium phosphate complexes of tryptic casein phosphopeptides in the rat. *J. Dent. Res.* 74, 1272–1279

62 Chabance, B. *et al.* (1998) Casein peptide release and passage to the blood in humans during digestion of milk or yogurt. *Biochimie* 80, 155–165

63 Meisel, H. *et al.* (2003) Detection of caseinophosphopeptides in the distal ileostomy fluid of human subjects. *Br. J. Nutr.* 89, 351–359

64 FitzGerald, R.J. *et al.* (2004) Hypotensive peptides from milk proteins. *J. Nutr.* 134, 980S–988S

65 Mullally, M.M. *et al.* (1997) Identification of a novel angiotensin-I-converting enzyme inhibitory peptide corresponding to a tryptic fragment of bovine beta-lactoglobulin. *FEBS Lett.* 402, 99–101

66 Robert, M.C. *et al.* (2004) Identification of angiotensin-I-converting enzyme inhibitory peptides derived from sodium caseinate hydrolysates produced by *Lactobacillus helveticus* NCC 2765. *J. Agric. Food Chem.* 52, 6923–6931

67 Baker, A.M. *et al.* (2005) Central penetration and stability of N-terminal tripeptide of insulin-like growth factor-I, glycine-proline-glutamate in adult rat. *Neuropeptides* 39, 81–87

68 Sizonenko, S.V. *et al.* (2001) Neuroprotective effects of the N-terminal tripeptide of IGF-1, glycine-proline-glutamate, in the immature rat brain after hypoxic-ischemic injury. *Brain Res.* 922, 42–50

69 Gupta, S.K. *et al.* (1995) A potent inhibitor of endothelial cell proliferation is generated by proteolytic cleavage of the chemokine platelet factor 4. *Proc. Natl. Acad. Sci. USA* 92, 7799–7803

70 Walsh, G. (2003) Biopharmaceutical benchmarks–2003. *Nat. Biotechnol.* 21, 865–870

71 Reichert, J.M. (2003) Trends in development and approval times for new therapeutics in the United States. *Nat. Rev. Drug Discov.* 2, 695–702

72 Heffernan, M. *et al.* (2001) The effects of human GH and its lipolytic fragment (AOD9604) on lipid metabolism following chronic treatment in obese mice and beta(3)-AR knock-out mice. *Endocrinology* 142, 5182–5189

73 Mizuno, S. *et al.* (2005) Antihypertensive effect of casein hydrolysate in a placebo-controlled study in subjects with high-normal blood pressure and mild hypertension. *Br. J. Nutr.* 94, 84–91

74 Iijima, Y. *et al.* (2004) Acid resistance of enamel subsurface lesions remineralized by a sugar-free chewing gum containing casein phosphopeptide-amorphous calcium phosphate. *Caries Res.* 38, 551–556

75 Ramalingam, L. *et al.* (2005) Adding casein phosphopeptide-amorphous calcium phosphate to sports drinks to eliminate *in vitro* erosion. *Pediatr. Dent.* 27, 61–67

76 Eder, J.P., Jr *et al.* (2002) Phase I clinical trial of recombinant human endostatin administered as a short intravenous infusion repeated daily. *J. Clin. Oncol.* 20, 3772–3784

77 Marshall, E. (2002) Cancer therapy. Setbacks for endostatin. *Science* 295, 2198–2199

78 Tjin Tham Sjin, R.M. *et al.* (2005) A 27-amino-acid synthetic peptide corresponding to the NH2-terminal zinc-binding domain of endostatin is responsible for its antitumor activity. *Cancer Res.* 65, 3656–3663

79 Subramanian, I.V. *et al.* (2005) Adeno-associated virus-mediated delivery of a mutant endostatin suppresses ovarian carcinoma growth in mice. *Gene Ther.* 12, 30–38

80 Yokoyama, Y. and Ramakrishnan, S. (2004) Improved biological activity of a mutant endostatin containing a single amino-acid substitution. *Br. J. Cancer* 90, 1627–1635

81 Daly, M.E. *et al.* (2003) Hemostatic regulators of tumor angiogenesis: a source of antiangiogenic agents for cancer treatment? *J. Natl. Cancer Inst.* 95, 1660–1673

82 Whitelock, J.M. *et al.* (1996) The degradation of human endothelial cell-derived perlecan and release of bound basic fibroblast growth factor by stromelysin, collagenase, plasmin, and heparanases. *J. Biol. Chem.* 271, 10079–10086

83 Zucht, H.D. *et al.* (1995) Casocidin-I: a casein-alpha s2 derived peptide exhibits antibacterial activity. *FEBS Lett.* 372, 185–188

Free journals for developing countries

The WHO and six medical journal publishers have launched the Access to Research Initiative, which enables nearly 70 of the world's poorest countries to gain free access to biomedical literature through the Internet.

Gro Harlem Brundtland, director-general for the WHO, said that this initiative was 'perhaps the biggest step ever taken towards reducing the health information gap between rich and poor countries'.

For more information, visit www.healthinternetwork.net