The Biological and Biomedical Consequences of Protein Moonlighting

A Biochemical Society Focused Meeting held at Charles Darwin House, London, U.K., 29–30 July 2014. Organized and Edited by Brian Henderson and Andrew Martin (University College London, U.K.).

Protein moonlighting: a new factor in biology and medicine

Brian Henderson* **and Andrew C. R. Martin†**

*Department of Microbial Diseases, UCL Eastman Dental Institute, 256 Gray's Inn Road, London WC1X 8LD, U.K.

†Biomolecular Structure and Modelling Unit, Institute of Structural and Molecular Biology, Research Department of Structural and Molecular Biology, Division of Biosciences, University College London, Gower Street, London WC1E 6BT, U.K.

Abstract

The phenomenon of protein moonlighting was discovered in the 1980s and 1990s, and the current definition of what constitutes a moonlighting protein was provided at the end of the 1990s. Since this time, several hundred moonlighting proteins have been identified in all three domains of life, and the rate of discovery is accelerating as the importance of protein moonlighting in biology and medicine becomes apparent. The recent re-evaluation of the number of protein-coding genes in the human genome (approximately 19000) is one reason for believing that protein moonlighting may be a more general phenomenon than the current number of moonlighting proteins would suggest, and preliminary studies of the proportion of proteins that moonlight would concur with this hypothesis. Protein moonlighting could be one way of explaining the seemingly small number of proteins that are encoded in the human genome. It is emerging that moonlighting proteins can exhibit novel biological functions, thus extending the range of the human functional proteome. The several hundred moonlighting proteins so far discovered play important roles in many aspects of biology. For example, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), heat-shock protein 60 (Hsp60) and tRNA synthetases play a wide range of biological roles in eukaryotic cells, and a growing number of eukaryotic moonlighting proteins are recognized to play important roles in physiological processes such as sperm capacitation, implantation, immune regulation in pregnancy, blood coagulation, vascular regeneration and control of inflammation. The dark side of protein moonlighting finds a range of moonlighting proteins playing roles in various human diseases including cancer, cardiovascular disease, HIV and cystic fibrosis. However, some moonlighting proteins are being tested for their therapeutic potential, including immunoglobulin heavy-chain-binding protein (BiP), for rheumatoid arthritis, and Hsp90 for wound healing. In addition, it has emerged over the last 20 years that a large number of bacterial moonlighting proteins play important roles in bacteria–host interactions as virulence factors and are therefore potential therapeutic targets in bacterial infections. So as we progress in the 21st Century, it is likely that moonlighting proteins will be seen to play an increasingly important role in biology and medicine. It is hoped that some of the major unanswered questions, such as the mechanism of evolution of protein moonlighting, the structural biology of moonlighting proteins and their role in the systems biology of cellular systems can be addressed during this period.

Key words: glyceraldehyde-3-phosphate dehydrogenase (GAPDH), heat-shock protein 60 (Hsp60), moonlighting.

Abbreviations: AMF, autocrine motility factor; BiP, immunoglobulin heavy-chain-binding protein; Cpn, chaperonin; ER, endoplasmic reticulum; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HER2, human epidermal growth factor receptor 2; Hsp, heat-shock protein; MAFMP, multiple additional function moonlighting protein; PGI, phosphoglucoisomerase; SLiM, short linear motif.

[1]**To whom correspondence should be addressed (email brian.henderson@ucl.ac.uk).**

Introduction

A key requirement in biology is to obtain a measure of the complexity of any particular organism. One obvious basis for this is the number of protein-coding genes an organism has at its disposal. Proteins are the essential tools of biology and it would be assumed that an organism as complex as the reader of the present article would need a 'large number' of such protein-coding genes. Thus before the human genome was sequenced, estimates of the number of protein-coding genes ranged from 40 000 to 100 000 and this was a re-estimate of earlier values that ranged up to 2 million genes [1]. Such numbers seemed reasonable for the maintenance of both the complex eukaryotic cells and for the incredible number of interactions that must exist between them to maintain the co-existence of the 10^{13} cells that compose the average human body. It therefore came as a surprise when the initial annotations of the human genome suggested that only 26 000–30 000 protein-coding genes existed [2,3]. Further annotation brought this number down to 20 000–25 000 protein-coding genes [4]. In the ensuing decade, these numbers have continued to fall, with the estimate of 20 500 being made in 2007 [5] and the most recent estimate suggesting the human genome contains only 19 000 protein-coding genes [6]. The human genome contains 3×10^9 bp which generate 19 000 proteins. Compare this with the water flea, *Daphnia pulex*, whose genome contains 0.2×10^9 bp, but encodes an estimated 31 000 genes [7]. This shows that the human genome could be much larger, but that it has evolved to have this relatively small number of genes as it is sufficient for purpose. Are there any other sources of complexity in the protein landscape? The human proteome is, of course, clearly more complex than the genome, due to the generation of splice and other protein variants. This is currently estimated to contain 10^6 proteoforms [8,9]. Alternative splicing of eukaryotic genes is recognized to be a common process [10], and can contribute to cellular complexity. However, although splice variants such as those of kinases show different expression levels in cell compartments, they retain similar functional characteristics [11], although it is now clear that they form part of a complex regulatory network [12]. So the question remains as to how the human organism can be put together with relatively few functional components.

One hypothesis which may explain this conundrum is that proteins are not monofunctional, but, depending on evolutionary constraints, can exhibit more than one unique biological activity. This can be termed the 'moonlighting hypothesis' and is the subject of the following papers in this issue of *Biochemical Society Transactions*. The definition of protein moonlighting being used in the present article is based on the classic review by Connie Jeffery [13].

A brief history of protein moonlighting

It is only in 2014 that individuals such as Juan Cedano [14] and Connie Jeffery [14a] are bringing together all known examples of moonlighting proteins into databases. These reveal that there are currently ~300 reported moonlighting proteins, although it is clear that a number of proteins are missing from the databases. However, these databases should be able to capture the data on the increasing numbers of moonlighting proteins that are constantly being discovered.

We can trace back the history of protein moonlighting to before the use of this term and there are a number of strands to this history. In one, the moonlighting activity being identified was the binding of metabolic proteins to ssDNA. This came from studies of transformed cells in which a DNA-binding protein, termed P8, was discovered. Synthesis of P8 was dependent on the cell cycle in normal cells, but this control was lost with transformation. A key finding was the ability of P8 to bind ssDNA [15]. However, it took a further 4 years to identify that P8 was the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) [16]. This same group showed that GAPDH also bound to RNA and that the DNA-binding capacity was inhibited by NAD^+, but not by the substrate, glyceraldehyde 3-phosphate, suggesting that binding of DNA was through the NAD^+-binding sites in the protein [17]. In later studies, it was shown that lactate dehydrogenase was also an ssDNA-binding protein [18].

Another strand came from an Australian group studying fetal implantation who hypothesized that implantation of the fetus, which is an immunologically foreign material, must require some form of immune regulation to limit anti-fetal immunity. This led to the discovery of early pregnancy factor (EPF) in 1977 [19,20], which was shown to be an immunosuppressant [21]. However, it was not until 1994 that this protein was identified as the molecular chaperone, chaperonin 10 (Cpn10) [22]. Just over a decade later, and this protein was in clinical trials as an immunosuppressant for treating rheumatoid arthritis [23].

In the 1980s, Graham Wistow and Joram Piatigorsky, working on the composition of the proteins of the lens of the eye, made the unexpected finding that some of these proteins were metabolic enzymes such as lactate dehydrogenase [24] or arginosuccinate lyase [25]. This led to the finding that many lens proteins are either metabolic enzymes or cell stress proteins, and Piatigorsky coined the term 'gene sharing' for this process [26,27]. The term did not catch on.

The late 1980s and early 1990s saw the emergence of publications on the secretion of what were thought to be purely intracellular proteins, i.e. molecular chaperones and protein-folding catalysts [28], and the finding that these proteins had potent cell signalling actions (reviewed in [29]). Thus human thioredoxin was initially discovered as a secreted cytokine from a T-cell lymphoma able to induce expression of the interleukin (IL)-2 receptor [30]. Similarly, the protein-folding catalyst cyclophilin A was shown to be secreted by lipopolysaccharide-stimulated monocytes and to be a potent pro-inflammatory protein with chemotactic activity [31]. Today, a number of these cell stress proteins are either key disease biomarkers [29] or are in clinical trial, for example immunoglobulin heavy-chain-binding protein (BiP) for rheumatoid arthritis (see Panayi and Corrigall in this issue

of *Biochemical Society Transactions* [32]) or, like heat-shock protein 90α (Hsp90α), are heading to clinical trial, in this case for wound repair [33].

Another strand concerns the glycolytic enzyme phosphoglucoisomerase (PGI), which, in the 1950s, had been found in the serum of cancer patients [34]. In the mid-1980s, a protein, then termed neuroleukin, was found which both enhanced lymphocyte immunoglobulin secretion [35] and was a neurotrophic factor for spinal sensory neurons [36]. Of note, this protein was active at concentrations between 10^{-9} and 10^{-11} M, implying high-affinity binding to a receptor. A few years later, neuroleukin was shown to be PGI [37]. Since then, PGI has been shown to have at least four other moonlighting actions, the most important of which is as autocrine motility factor (AMF), a secreted autocrine factor which promotes epithelial–mesenchymal transformation, important in carcinogenesis. This was shown to be PGI in the late 1990s [38].

The 1990s also revealed that moonlighting proteins played a role in the communication between both bacterial and protozoal pathogens and the human host. The first evidence for this was that the glycolytic enzyme GAPDH was able to act as an adhesin, invasin, toxin and evasin for the major bacterial pathogen *Streptococcus pyogenes*. These are the four main classes of bacterial virulence factors, and for one protein to exhibit all such functions is, if not unique, then very rare (reviewed in [39]). The role of moonlighting proteins in bacteria and protozoal infections is covered in this issue of *Biochemical Society Transactions* by Henderson [40] and Ginger [41] respectively.

In 1999, Connie Jeffery wrote her classic review [13] of the literature on proteins with more than one function and coined the term protein moonlighting and codified what a moonlighting protein was. This definition is still the basis of our understanding of what a moonlighting protein is, although there is clearly scope for modifying the definition. In the 15 years since the publication of her review, and because her review had made protein moonlighting a focus of research, a growing number of moonlighting proteins have emerged.

Moonlighting proteins and the protein interactome

Proteins function within a complex network of interactions, both with other proteins and with non-proteins, that now has its own '-omics' terminology, i.e. the interactome, and this is now clearly seen to be a key factor in human diseases [42]. The interactome is governed by two factors: the affinity of interaction and the number of interactions for each protein. One criticism that is often made of moonlighting proteins is that their moonlighting actions are only the result of non-specific binding events, because they bind to their assumed target with low affinities. If this is true, it clearly weakens the case for those proposing that moonlighting proteins contribute to the function of the cellular interactome and in so doing have important biological functions. Protein ligand-binding affinities (K_d values) normally range from $<10^{-6}$ M to $>10^{12}$ M with K_d values of $<10^{-8}$ M being considered to be of high affinity [43]. A range of moonlighting proteins have the capacity to bind to human plasmin(ogen) and, where the affinity has been measured, it is generally in the nanomolar range. Over the years, a number of bacterial (and some human) moonlighting proteins have been assessed for their binding affinities (Table 1). Clearly, it can be seen that these various moonlighting proteins bind with high affinity to their respective ligands. This supports the hypothesis that protein moonlighting interactions within the cellular and extracellular interactome are biologically important.

Protein moonlighting in the three domains of life

Carl Woese proposed that all organisms on the planet can fit into three cellular domains: Bacteria, Archaea and Eukarya [44]. Can we gain any information from asking the question about the evolution of protein moonlighting in these three evolutionary compartments which have had distinct evolutionary trajectories. This question, of course, will be influenced by the research effort focused on these distinct domains. At the time of writing, we have identified only five moonlighting proteins in the Archaea [45]. As in the other domains, these proteins are either molecular chaperones or metabolic enzymes. In contrast, there are approximately 100 bacterial moonlighting proteins and >100 eukaryotic moonlighting proteins. Of interest, many of these moonlighting proteins have homologues in both the Bacteria and the Eukarya (and possibly also in the Archaea). So it is not possible to determine whether evolution has favoured one cellular domain of life over another, as the differences may come down to the range of proteins being examined by scientists studying the three basic cell types. The only thing that can be said is that protein moonlighting exists in all three domains of life.

Evolution of protein moonlighting and its role in the molecular toolkit

Understanding the evolution of protein moonlighting has to encompass some curious findings. Thus, at the time of writing, moonlighting proteins can be considered to fall into two categories: (i) single additional function moonlighting proteins (SAFMPs), and (ii) multiple additional function moonlighting proteins (MAFMPs). With some of the latter, such as GAPDH [46] or the molecular chaperone Cpn60 [47], these protein families can exhibit between 20 and 40 distinct biological functions. With GAPDH, the human enzyme itself has approximately 15–20 different cellular actions [46]. It is also interesting that most of the MAFMPs are evolutionarily ancient proteins including: Cpn10, enolase, dihydrolipoamide dehydrogenase, aconitase, Hsp70 and EF-Tu. Protein evolution is still something of a controversial

Table 1 | Binding affinities of moonlighting proteins to ligands other than plasmin(ogen)
The references to these individual studies can be found in [45].

Species	Moonlighting protein	Ligand bound	K_d (M)
Mycobacterium tuberculosis	GAPDH	Epidermal growth factor	2×10^{-10}
Listeria monocytogenes	Alcohol acetaldehyde dehydrogenase	Human Hsp60	5.4×10^{-8}
Pasteurellaceae	ComE1 proteins	Fibronectin	150×10^{-9}
Yersinia pestis	Caf1	IL-1 receptor	5.4×10^{-10}
Yersinia pestis	Caf1A usher protein	IL-1β	1.4×10^{-10}
Yersinia pestis	Lcrv	IFNγ	32×10^{-9}
Streptococcus pyogenes	Protein H	Fibronectin	16×10^{-8}
Streptococcus pyogenes	Serum opacity factor 2	Fibulin-1	1.6×10^{-9}
Streptococcus pneumoniae	GAPDH	C1q	3×10^{-10}
Homo sapiens	GAPDH	C1q	2.9×10^{-9}
Homo sapiens	PGI	GP78	$0.01\text{–}1\times10^{-9}$

arena with a number of conflicting theories about how protein function evolves [48]. Mutations are constantly occurring in genes, giving rise to the concept of the molecular clock, and one of the major hypotheses to account for protein evolution suggested the novel and controversial concept that among mutations that go to fixation, the vast majority are selectively neutral. This is Kimura's neutral theory of molecular evolution [49], and it is has to be assumed that it is among these neutral mutations that moonlighting sites evolve. Other chapters in this issue of *Biochemical Society Transactions* by Shelley Copley [50], Mario Fares [51] and Carlos Gancedo [52] focus on the potential evolutionary mechanisms for the development of protein moonlighting sites. However, we want to focus on some concrete information about these sites. In the Cpn60 protein family, a moderate number of distinct sites have been identified to account for individual biological actions. These all appear to be found on the protein surface and appear to range in size from single residues to >50 residues [47]. In the case of GAPDH, only a few sites have been ascribed a function [46], but there is a small amount of literature suggesting that protein modification of GAPDH is associated with the protein exhibiting moonlighting functions. For example, deamidation is required for the protein to exhibit fusogenic activity [53].

One of the most confusing aspects of moonlighting proteins, which is relevant for this discussion, is the finding that protein 'homologues' and 'paralogues' can exhibit different biological activities in spite of high levels of sequence identity. The two *Mycobacterium tuberculosis* Cpn60 paralogues (>60% sequence identity) have completely different actions on bone. The Cpn60.1 protein inhibits the formation of osteoclasts, whereas the Cpn60.2 protein has no activity at all, and the *Escherichia coli* Cpn60 protein is a highly potent stimulator of osteoclast formation [54]. Similarly the EF-Tu proteins of *Mycoplasma pneumoniae* and *Mycoplasma genitalium*, which have 96% sequence identity, have completely different binding interactions with fibronectin [55]. This implies that these moonlighting sites contain only a few residues. Short protein motifs, as short as three residues [56], appear to be the reason for the diversification of protein isoforms (proteoforms) described earlier [57]. These are known as short linear motifs (SLiMs) and often associate with intrinsically disordered regions in proteins, which itself has been claimed to be associated with protein moonlighting [58]. It has recently been estimated that there are 10^6 SLiMs in the human proteome [59], providing an enormous number of potential moonlighting sites in humans. So are the moonlighting sites in proteins SLiMs? The only objection to this idea is that SLiMs generally exhibit low-affinity binding, whereas, where measured (Table 1), virtually all moonlighting proteins have K_d values in the subnanomolar range. Further evidence that moonlighting sites contain only small numbers of residues comes from Andrew Martin's analysis [60] (in this issue of *Biochemical Society Transactions*) of the antibody-combining site, in which he has found that changes of only a few residues can markedly alter antigenic specificity.

If protein moonlighting is an evolutionary solution allowing organisms to limit the number of protein-coding genes they have to carry, then one might expect that most of these moonlighting functions would be novel. There are a number of examples of moonlighting proteins having unexplained, or novel, functions. Thus the secreted mouse Hsp90α protein is a novel epithelial cell-modulating protein with wound-healing properties [33]. The *Legionella pneumophila* Cpn60 protein, when inside human cells, functions as a mitochondrial chemoattractant [61]. Finally, PGI regulates the endoplasmic reticulum (ER) stress response by a novel calcium-dependent process [62]. A number of other examples of moonlighting proteins with novel actions are known, and some are described below. It is clear that this search for novel protein moonlighting functions must be a key focus of future research into protein moonlighting and the question of whether protein moonlighting significantly extends the 'molecular toolkit' of the cell can only really be addressed when many more such proteins have been identified.

Protein moonlighting in cell biology

As more moonlighting proteins are identified, their roles in the functioning of eukaryotic and prokaryotic cells slowly start to emerge. These moonlighting actions are generally unexpected as with the role of mitochondrial cytochrome *c* in apoptosis [63], the cytosolic transcription factor signal transducer and activator of transcription 3 (STAT3) in the mitochondrial electron transport chain [64] or the glycolytic enzyme phosphofructokinase in autophagy [65].

There is a general unstated belief that, although protein moonlighting may actually be a real phenomenon, the moonlighting actions of any particular protein is subservient to its 'true' activity. Thus the glycolytic activity of glycolytic enzymes would be assumed to be more important than their various moonlighting actions. This is generally difficult to test as glycolytic enzymes are essential and their genes cannot be inactivated. In *Neisseria*, the glycolytic pathway does not function, as the gene for phosphofructokinase is lacking [66]. In spite of this, both the aldolase [67] and enolase enzymes [68] are defined virulence factors in *Neisseria meningitidis*. This is also the conclusion of a recent review on the multiple moonlighting actions of tRNA synthetases, with the moonlighting actions of these proteins being as critical for cellular homoeostasis as their activity in protein translation [69]. This starts to raise the chicken-and-egg question about protein moonlighting in terms of what really is the moonlighting function of mutlifunctional proteins.

The realization that proteins can exhibit more than one biological activity immediately begins to increase the complexity of the interactome of each moonlighting protein. How much the existence of protein moonlighting will influence the interactomic complexity within cells, and also cause a similar increase in complexity in transcriptional and translational activity, will depend on what proportion of the 19000 proteins in *Homo sapiens*, for example, moonlight. As far as we are aware, only two experimental studies shed light on the commonality of protein moonlighting. In a study of DNA-binding proteins, over 4000 non-redundant human proteins were expressed, with over half not being expected to bind DNA. However, 22.4% of these non-DNA-binding proteins actually bound to DNA [70]. A separate study screened a metagenomic library of oral bacteria in filamentous phage M13 against three protein ligands: fibronectin, IgA and BSA. Surprisingly, 30–40% of the positive clones contained genes for known metabolic enzymes, molecular chaperones or transporter proteins (B. Henderson, unpublished work). Studies such as these suggest that a substantial proportion of cellular proteins may actually be moonlighting proteins and the inference is that cellular systems are much more complex than we currently imagine, due to the participation of moonlighting proteins.

Medical aspects of protein moonlighting

Moonlighting proteins are involved in homoeostatic regulation in the human and in human reproduction. The role of Cpn10 in implantation has been discussed and PGI is also reported to be an implantation factor in the ferret [71]. The role of moonlighting proteins in the interaction of sperm and egg is detailed in this issue of *Biochemical Society Transactions* by Francois Petit [72]. Other moonlighting proteins involved in normal physiology include mammalian GAPDH which is a novel autocrine iron-uptake protein binding to transferrin [73], thymidine phosphorylase, initially identified as platelet-derived growth factor (PDGF), which has angiogenic and chemotactic actions [74], and protein disulfide-isomerases, molecular chaperones which are also involved in platelet aggregation and controlling blood clotting [75]. Another fascinating moonlighting protein with a key homoeostatic function is the ubiquitin analogue ISG15. Note that ubiquitin is also a moonlighting peptide with multiple functions [76]. It has recently emerged that ISG15 is a natural secreted control factor for the production of interferon γ (IFNγ) which has the consequence that it maintains our immunity to mycobacterial infections. Given that it is estimated that one-third of the world's population are latently infected with the causative agent of tuberculosis, *Mycobacterium tuberculosis*, and that this is a chronic inflammatory disease, it is clear that ISG15 is a very important factor in the control of human inflammation. Individuals lacking ISG15 are very susceptible to mycobacterial infection [77].

In the future, the study of protein moonlighting will focus on its role in medicine, as the number of moonlighting proteins acting either as pathological factors, or indeed, as therapeutic agents, increases. We return to a consideration of the glycolytic enzyme, PGI, as the pro-cancer factor, AMF [38]. This protein induces epithelial–mesenchymal transformation, a process in which epithelial cells loose cell polarity and cell–cell adhesion and become migratory and invasive, potentially leading to metastasis, and circulating levels of PGI correlate with poor prognosis in breast cancer [78]. The human epidermal growth factor receptor 2 (HER2) is a therapeutic target in breast cancer which is blocked by the human monoclonal antibody Herceptin®. PGI binds to HER2 and is able to induce its proteolytic removal thus rendering such cancers refractory to Herceptin® [79].

The vasculature is the scene of action of certain mitochondrial proteins. The mitochondrial ATP synthase is a mitochondrial complex consisting of three domains, namely the extrinsic and intrinsic membrane domains (F_1 and F_0 respectively) joined by a stalk. Four subunits of the stalk have been designated coupling factor 6, oligomycin sensitivity-conferring protein, and subunits b and d. Mitochondrial coupling factor 6 was reported to be essential for energy transduction [80]. There is now strong evidence that this protein plays important roles in cardiovascular disease [81] by interacting with a plasma membrane-bound ATP synthase which is itself a moonlighting protein involved in human disease pathology [82]. As space is limited, a number of other examples of moonlighting proteins contributing to virulence is shown in Table 2.

In addition to moonlighting proteins playing a role in idiopathic diseases, there is now strong evidence that bacterial

Table 2 | **Some examples of moonlighting proteins involved in human disease states**

Moonlighting protein	Disease state or pathological activity	Reference(s)
PGI	Human cancer	[38,78,79]
Mitochondrial coupling factor 6	Vascular disease	[81]
ATP synthase	Vascular disease/vasculitides	[82]
Cyclophilin A	Atherosclerosis and other diseases	[83]
Thioredoxin	HIV	[84]
High-mobility group box 1 (HMGB1)	Various disease states	[85]
Cystic fibrosis transmembrane conductance regulator (CFTR)	Cystic fibrosis	[86]
Enolase	Cancer	[87]
Succinate dehydrogenase	Tumorigenesis	[88]
Fumarate hydratase	Tumorigenesis	[89]
Phosphoglycerate kinase	Metastasis promoter	[90]
Hsp90	Epithelial–mesenchymal transformation (like PGI)	[91]

pathogens, interacting with the human organism, use a multitude of moonlighting proteins as virulence factors (see the article by Henderson [40] in this issue of *Biochemical Society Transactions*).

On the other side of this medical coin, it is emerging that moonlighting proteins have therapeutic potential. The chapter by Panayi and Corrigall [32] reveals that the ER molecular chaperone BiP may have unique therapeutic properties, which might lead to a cure for rheumatoid arthritis.

Conclusions

Protein moonlighting, as a scientific discipline, is in its infancy and it is certain that by the time the next conference on this topic is held, many more moonlighting proteins will have been identified and it is only with an increased population of such proteins that we can begin to develop hypotheses as to the role moonlighting proteins play in biology and medicine. Obvious questions to address are: (i) what is the mechanism of the evolution of moonlighting proteins?; (ii) what is the structural basis of protein moonlighting?; (iii) does protein moonlighting add to the systems complexity of the cell?; (iv) what is the role of protein moonlighting in human genetics?; (v) what roles do moonlighting proteins play in human disease?; (vi) how many moonlighting proteins are likely to be therapeutic agents?; and (vii) does protein moonlighting contribute to adverse drug responses? It is hoped that answers to these questions can be reached within the next decade.

Funding

We are grateful to the Biotechnology and Biological Sciences Research Council, the Medical Research Council and the Wellcome Trust for financial support.

References

1 Pennisi, E. (2003) A low gene number wins the GeneSweep pool. Science **300**, 1484 CrossRef PubMed
2 Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. et al. (2001) The sequence of the human genome. Science **291**, 1304–1351 CrossRef PubMed
3 Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. et al. (2001) Initial sequencing and analysis of the human genome. Nature **409**, 860–921 CrossRef PubMed
4 Finishing the euchromatic sequence of the human genome (2004) Nature **431**, 931–945, International Human Genome Sequencing Consortium CrossRef PubMed
5 Clamp, M., Fry, B., Kamal, M., Xie, X., Cuff, J., Lin, M.F., Kellis, M., Lindblad-Toh, K. and Lander, E.S. (2007) Distinguishing protein-coding and noncoding genes in the human genome. Proc. Natl. Acad. Sci. U.S.A. **104**, 19428–19433 CrossRef PubMed
6 Ezkurdia, I., Juan, D., Rodriguez, J.M., Frankish, A., Diekhans, M., Harrow, J., Vazquez, J., Valencia, A. and Tress, M.L. (2014) Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. Hum. Mol. Genet. **23**, 5866–5878 PubMed
7 Colbourne, J.K., Pfrender, M.E., Gilbert, D., Thomas, W.K., Tucker, A., Oakley, T.H., Tokishita, S., Aerts, A., Arnold, G.J., Basu, M.K. et al. (2011) The ecoresponsive genome of *Daphnia pulex*. Science **331**, 555–561 CrossRef PubMed
8 Muñoz, J. and Heck, A.J. (2014) From the human genome to the human proteome. Angew. Chem. Int. Ed. Engl. **53**, 10864–10866 CrossRef PubMed
9 Smith, L.M. and Kelleher, N.L. (2013) Proteoform: a single term describing protein complexity. Nat. Methods **10**, 186–187 CrossRef PubMed
10 Pal, S., Gupta, R. and Davuluri, R.V. (2012) Alternative transcription and alternative splicing in cancer. Pharmacol. Ther. **136**, 283–294 CrossRef PubMed
11 Anamika, K., Garnier, N. and Srinivasan, N. (2009) Functional diversity of human protein kinase splice variants marks significant expansion of human kinome. BMC Genomics **10**, 622 CrossRef PubMed
12 Calarco, J.A., Zhen, M. and Blencowe, B.J. (2011) Networking in a global world: establishing functional connections between neural splicing regulators and their target transcripts. RNA **17**, 775–791 CrossRef PubMed
13 Jeffery, C.J. (1999) Moonlighting proteins. Trends Biochem. Sci. **24**, 8–11 CrossRef PubMed
14 Hernández, S., Ferragut, G., Amela, I., Perez-Pons, J., Piñol, J., Mozo-Villarias, A., Cedano, J. and Querol, E. (2013) MultitaskProtDB: a database of multitasking proteins. Nucleic Acids Res. **42**, D517–D520 PubMed
14a Mani, M., Chen, C., Amblee, V., Liu, H., Mathur, T., Zwicke, G., Zabad, S., Patel, B., Thakkar, J. and Jeffery, C. J. (2014) MoonProt: a database for proteins that are known to moonlight. Nucleic Acids Res., doi:10.1043/nar/gku954 PubMed

15 Tsai, R.L. and Green, H. (1973) Studies on a mammalian cell protein (P8) with affinity for DNA *in vitro*. J. Mol. Biol. **73**, 307–316 CrossRef PubMed

16 Perucho, M., Salas, J. and Salas, M.L. (1977) Identification of the mammalian DNA-binding protein P8 as glyceraldehyde-3-phosphate dehydrogenase. Eur. J. Biochem. **81**, 557–562 CrossRef PubMed

17 Perucho, M., Salas, J. and Salas, M.L. (1980) Study of the interaction of glyceraldehyde-3-phosphate dehydrogenase with DNA. Biochim. Biophys. Acta **606**, 181–195 CrossRef PubMed

18 Williams, K.R., Reddigari, S. and Patel, G.L. (1985) Identification of a nucleic acid helix-destabilizing protein from rat liver as lactate dehydrogenase-5. Proc. Natl. Acad. Sci. U.S.A. **82**, 5260–5264 CrossRef PubMed

19 Morton, H., Rolfe, B. and Clunie, G.J. (1977) An early pregnancy factor detected in human serum by the rosette inhibition test. Lancet **1**, 394–397 CrossRef PubMed

20 Morton, H. (1998) Early pregnancy factor: an extracellular chaperonin 10 homologue. Immun. Cell Biol. **76**, 483–496 CrossRef

21 Noonan, F.P., Halliday, W.J., Morton, H. and Clunie, G.J. (1979) Early pregnancy factor is immunosuppressive. Nature **278**, 649–651 CrossRef PubMed

22 Cavanagh, A.C. and Morton, H. (1994) The purification of early-pregnancy factor to homogeneity from human platelets and identification as chaperonin 10. Eur. J. Biochem. **222**, 551–560 CrossRef PubMed

23 Vanags, D., Williams, B., Johnson, B., Hall, S., Nash, P., Taylor, A., Weiss, J. and Feeney, D. (2006) Therapeutic efficacy and safety of chaperonin 10 in patients with rheumatoid arthritis: a double-blind randomised trial. Lancet **368**, 855–863 CrossRef PubMed

24 Wistow, G.J., Mulders, J.W. and de Jong, W.W. (1987) The enzyme lactate dehydrogenase as a structural protein in avian and crocodilian lenses. Nature **326**, 622–624 CrossRef PubMed

25 Piatigorsky, J., O'Brien, W.E., Norman, B.L., Kalumuck, K., Wistow, G.J., Borras, T., Nickerson, J.M. and Wawrousek, E.F. (1988) Gene sharing by δ-crystallin and argininosuccinate lyase. Proc. Natl. Acad. Sci. U.S.A. **85**, 3479–3483 CrossRef PubMed

26 Piatigorsky, J. and Wistow, G.J. (1989) Enzyme/crystallins: gene sharing as an evolutionary strategy. Cell **57**, 197–199 CrossRef PubMed

27 Piatigorsky, J. (2007) Gene Sharing and Protein Evolution, Harvard University Press, Cambridge, MA., U.S.A.

28 Tytell, M., Greenberg, S.G. and Lasek, R.J. (1986) Heat shock-like protein is transferred from glia to axon. Brain Res. **363**, 161–164 CrossRef PubMed

29 Henderson, B. and Pockley, A.G. (2012) Proteotoxic stress and circulating cell stress proteins in the cardiovascular diseases. Cell Stress Chaperones **17**, 303–311 CrossRef PubMed

30 Tagaya, Y., Maeda, Y., Mitsui, A., Kondo, N., Matsui, H., Hamuro, J., Brown, N., Arai, K., Yokota, T., Wakasugi, H. et al. (1989) ATL-derived factor (ADF), an IL-2 receptor/Tac inducer homologous to thioredoxin; possible involvement of dithiol-reduction in the IL-2 receptor induction. EMBO J. **8**, 757–764 PubMed

31 Sherry, B., Yarlett, N., Strupp, A. and Cerami, A. (1992) Identification of cyclophilin as a proinflammatory secretory product of lipopolysaccharide-activated macrophages. Proc. Natl. Acad. Sci. U.S.A. **89**, 3511–3515 CrossRef PubMed

32 Panayi, G.S. and Corrigall, V.M. (2014) Immunoglobulin heavy-chain-binding protein (BiP): a stress protein that has the potential to be a novel therapy for rheumatoid arthritis. Biochem. Soc. Trans. **42**, 1752–1755

33 Tsen, F., Bhatia, A., O'Brien, K., Cheng, C.F., Chen, M., Hay, N., Stiles, B., Woodley, D.T. and Li, W. (2013) Extracellular heat shock protein 90 signals through subdomain II and the NPVY motif of LRP-1 receptor to Akt1 and Akt2: a circuit essential for promoting skin cell migration *in vitro* and wound healing *in vivo*. Mol. Cell. Biol. **33**, 4947–4959 CrossRef PubMed

34 Rossotto, P. and Ruffo, A. (1958) Prognostic significance of the study of certain enzymes of anaerobic glycolysis (aldolase, phosphoglucoisomerase) in the blood of malignant tumor patients. Minerva Med. **49**, 2145–2150 PubMed

35 Gurney, M.E., Apatoff, B.R., Spear, G.T., Baumel, M.J., Antel, J.P., Bania, M.B. and Reder, A.T. (1986) Neuroleukin: a lymphokine product of lectin-stimulated T cells. Science **234**, 574–581 CrossRef PubMed

36 Gurney, M.E., Heinrich, S.P., Lee, M.R. and Yin, H.S. (1986) Molecular cloning and expression of neuroleukin, a neurotrophic factor for spinal and sensory neurons. Science **234**, 566–574 CrossRef PubMed

37 Chaput, M., Claes, V., Portetelle, D., Cludts, I., Cravador, A., Burny, A., Gras, H. and Tartar, A. (1988) The neurotrophic factor neuroleukin is 90 % homologous with phosphohexose isomerase. Nature **332**, 454–455 CrossRef PubMed

38 Niinaka, Y., Paku, S., Haga, A., Watanabe, H. and Raz, A. (1998) Expression and secretion of neuroleukin/phosphohexose isomerase/maturation factor as autocrine motility factor by tumor cells. Cancer Res. **58**, 2667–2674 PubMed

39 Jin, H., Agarwal, S., Agarwal, S. and Pancholi, V. (2011) Surface export of GAPDH/SDH, a glycolytic enzyme, is essential for *Streptococcus pyogenes* virulence. MBio **2**, e00068–11 CrossRef PubMed

40 Henderson, B. (2014) An overview of protein moonlighting in bacterial infection. Biochem. Soc. Trans. **42**, 1720–1727

41 Ginger, M.L. (2014) Protein moonlighting in parasitic protists. Biochem. Soc. Trans. **42**, 1734–1739

42 Lage, K. (2014) Protein–protein interactions and genetic diseases: the interactome. Biochim. Biophys. Acta **1842**, 1971–1980 CrossRef PubMed

43 Houk, K.N., Leach, A.G., Kim, S.P. and Zhang, X. (2003) Binding affinities of host–guest, protein–ligand, and protein–transition-state complexes. Angew. Chem. Int. Ed. Engl. **42**, 4872–4897 CrossRef PubMed

44 Woese, C.R., Kandler, O. and Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. U.S.A. **87**, 4576–4579 CrossRef PubMed

45 Henderson, B., Fares, M. and Martin, A. (2015) Moonlighting Proteins in Biology and Medicine, Wiley-Blackwell, Chichester, in the press

46 Sirover, M.A. (2011) On the functional diversity of glyceraldehyde-3-phosphate dehydrogenase: biochemical mechanisms and regulatory control. Biochim. Biophys. Acta **1810**, 741–751 CrossRef PubMed

47 Henderson, B., Fares, M.A. and Lund, P.A. (2013) Chaperonin 60: a paradoxical, evolutionarily conserved protein family with multiple moonlighting functions. Biol. Rev. Camb. Philos. Soc. **88**, 955–987 PubMed

48 Akashi, H., Osada, N. and Ohta, T. (2012) Weak selection and protein evolution. Genetics **19**, 15–31 CrossRef

49 Kimura, M. (1985) The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge

50 Copley, S.D. (2014) An evolutionary perspective on protein moonlighting. Biochem. Soc. Trans. **42**, 1684–1691

51 Fares, M.A. (2014) The evolution of protein moonlighting: adaptive traps and promiscuity in the chaperonins. Biochem. Soc. Trans. **42**, 1709–1714

52 Gancedo, C., Flores, C.-L. and Gancedo, J.M. (2014) Evolution of moonlighting proteins: insight from yeasts. Biochem. Soc. Trans. **42**, 1715–1719

53 Iwai, K., Shibukawa, Y., Yamazaki, N. and Wada, Y. (2014) Transglutaminase 2-dependent deamidation of glyceraldehyde-3-phosphate dehydrogenase promotes trophoblastic cell fusion. J. Biol. Chem. **289**, 4989–4999 CrossRef PubMed

54 Winrow, V.R., Mesher, J., Meghji, S., Morris, C.J., Maguire, M., Fox, S., Coates, A.R., Tormay, P., Blake, D.R. and Henderson, B. (2008) The two homologous chaperonin 60 proteins of *Mycobacterium tuberculosis* have distinct effects on monocyte differentiation into osteoclasts. Cell. Microbiol. **10**, 2091–2104 CrossRef PubMed

55 Balasubramanian, S., Kannan, T.R., Hart, P.J. and Baseman, J.B. (2009) Amino acid changes in elongation factor Tu of *Mycoplasma pneumoniae* and *Mycoplasma genitalium* influence fibronectin binding. Infect. Immun. **77**, 3533–3541 CrossRef PubMed

56 Kubrycht, J., Sigler, K, Souček, P. and Hudeček, J. (2013) Structures composing protein domains. Biochimie **95**, 1511–1524 CrossRef PubMed

57 Weatheritt, R.J. and Gibson, T.J. (2012) Linear motifs: lost in (pre)translation. Trends Biochem. Sci. **37**, 333–341 CrossRef PubMed

58 Tompa, P., Szász, C. and Buday, L. (2005) Structural disorder throws new light on moonlighting. Trends Biochem. Sci. **30**, 484–489 CrossRef PubMed

59 Tompa, P., Davey, N.E., Gibson, T.J. and Babu, M.M. (2014) A million peptide motifs for the molecular biologist. Mol. Cell **55**, 161–169 CrossRef PubMed

60 Martin, A.C.R. (2014) Structural biology of moonlighting: lessons from antibodies. Biochem. Soc. Trans. **42**, 1704–1708

61 Chong, A., Lima, C.A., Allan, D.S., Nasrallah, G.K. and Garduno, R.A. (2009) The purified and recombinant *Legionella pneumophila* chaperonin alters mitochondrial trafficking and microfilament organization. Infect. Immun **77**, 4724–4739 CrossRef PubMed

62 Fu, M., Li, L., Albrecht, T., Johnson, J.D., Kojic, L.D. and Nabi, I.R. (2011) Autocrine motility factor/phosphoglucose isomerase regulates ER stress and cell death through control of ER calcium release. Cell Death Differ. **18**, 1057–1070 CrossRef PubMed

63 Hao, Z., Duncan, G.S., Chang, C.C., Elia, A., Fang, M., Wakeham, A., Okada, H., Calzascia, T., Jang, Y., You-Ten, A. et al. (2005) Specific ablation of the apoptotic functions of cytochrome *c* reveals a differential requirement for cytochrome *c* and Apaf-1 in apoptosis. Cell **121**, 579–591 CrossRef PubMed

64 Wegrzyn, J., Potla, R., Chwae, Y.J., Sepuri, N.B., Zhang, Q., Koeck, T., Derecka, M., Szczepanek, K., Szelag, M., Gornicka, A. et al. (2009) Function of mitochondrial Stat3 in cellular respiration. Science **323**, 793–797 CrossRef PubMed

65 Yuan, W., Tuttle, D.L., Shi, Y.J., Ralph, G.S. and Dunn, W.A. (1997) Glucose-induced microautophagy in *Pichia pastoris* requires the α-subunit of phosphofructokinase. J. Cell Sci. **110**, 1935–1945 PubMed

66 Baart, G.J., Zomer, B., de Haan, A., van der Pol, L.A., Beuvery, E.C., Tramper, J. and Martens, D.E. (2007) Modeling *Neisseria meningitidis* metabolism: from genome to metabolic fluxes. Genome Biol. **8**, R136 CrossRef PubMed

67 Tunio, S.A., Oldfield, N.J., Berry, A., Ala'Aldeen, D.A., Wooldridge, K.G. and Turner, D.P. (2010) The moonlighting protein fructose-1,6-bisphosphate aldolase of *Neisseria meningitidis*: surface localization and role in host cell adhesion. Mol. Microbiol. **76**, 605–615 CrossRef PubMed

68 Knaust, A., Weber, M.V., Hammerschmidt, S., Bergmann, S., Frosch, M. and Kurzai, O. (2007) Cytosolic proteins contribute to surface plasminogen recruitment of *Neisseria meningitidis*. J. Bacteriol. **189**, 3246–3255 CrossRef PubMed

69 Guo, M. and Schimmel, P. (2013) Essential nontranslational functions of tRNA synthetases. Nat. Chem. Biol. **9**, 145–153 CrossRef PubMed

70 Hu, S., Xie, Z., Onishi, A., Yu, X., Jiang, L., Lin, J., Rho, H.S., Woodard, C., Wang, H., Jeong, J.S. et al. (2009) Profiling the human protein–DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. Cell **139**, 610–622 CrossRef PubMed

71 Schulz, L.C. and Bahr, J.M. (2003) Glucose-6-phosphate isomerase is necessary for embryo implantation in the domestic ferret. Proc. Natl. Acad. Sci. U.S.A. **100**, 8561–8566 CrossRef PubMed

72 Petit, F.M., Serres, C. and Auer, J. (2014) Moonlighting proteins in sperm–egg interactions. Biochem. Soc. Trans. **42**, 1740–1743

73 Sheokand, N., Kumar, S., Malhotra, H., Tillu, V., Raje, C.I. and Raje, M. (2013) Secreted glyceraldehye-3-phosphate dehydrogenase is a multifunctional autocrine transferrin receptor for cellular iron acquisition. Biochim. Biophys. Acta **1830**, 3816–3827 CrossRef PubMed

74 Brown, N.S. and Bicknell, R. (1998) Thymidine phosphorylase, 2-deoxy-D-ribose and angiogenesis. Biochem. J. **334**, 1–8 PubMed

75 Wang, L., Wu, Y., Zhou, J., Ahmad, S.S., Mutus, B., Garbi, N., Hämmerling, G., Liu, J. and Essex, D.W. (2013) Platelet-derived ERp57 mediates platelet incorporation into a growing thrombus by regulation of the αIIbβ3 integrin. Blood **122**, 3642–3650 CrossRef PubMed

76 Majetschak, M. (2011) Extracellular ubiquitin: immune modulator and endogenous opponent of damage-associated molecular pattern molecules. J. Leukoc. Biol. **89**, 205–219 CrossRef PubMed

77 Bogunovic, D., Byun, M., Durfee, L.A., Abhyankar, A., Sanal, O., Mansouri, D., Salem, S., Radovanovic, I., Grant, A.V., Adimi, P. et al. (2012) Mycobacterial disease and impaired IFN-γ immunity in humans with inherited ISG15 deficiency. Science **337**, 1684–1688 CrossRef PubMed

78 Ahmad, A., Aboukameel, A., Kong, D., Wang, Z., Sethi, S., Chen, W., Sarkar, F.H. and Raz, A. (2011) Phosphoglucose isomerase/autocrine motility factor mediates epithelial–mesenchymal transition regulated by miR-200 in breast cancer cells. Cancer Res. **71**, 3400–3409 CrossRef PubMed

79 Kho, D.H., Nangia-Makker, P., Balan, V., Hogan, V., Tait, L., Wang, Y. and Raz, A. (2013) Autocrine motility factor promotes HER2 cleavage and signaling in breast cancer cells. Cancer Res. **73**, 1411–1419 CrossRef PubMed

80 Knowles, A.F., Guillory, R.J. and Racker, E. (1971) Partial resolution of the enzymes catalyzing oxidative phosphorylation. XXIV. A factor required for the binding of mitochondrial adenosine triphosphatase to the inner mitochondrial membrane. J. Biol. Chem. **246**, 2672–2679 PubMed

81 Osanai, T., Magota, K. and Okumura, K. (2009) Coupling factor 6 as a novel vasoactive and proatherogenic peptide in vascular endothelial cells. Naunyn Schmiedebergs Arch. Pharmacol. **380**, 205–214 CrossRef PubMed

82 Fu, Y. and Zhu, Y. (2010) Ectopic ATP synthase in endothelial cells: a novel cardiovascular therapeutic target. Curr. Pharm. Des. **16**, 4074–4079 CrossRef PubMed

83 Nigro, P., Pompilio, G. and Capogrossi, M.C. (2013) Cyclophilin A: a key player for human disease. Cell Death Dis. **4**, e888 CrossRef PubMed

84 Nakamura, H., De Rosa, S.C., Yodoi, J., Holmgren, A., Ghezzi, P., Herzenberg, L.A. and Herzenberg, L.A. (2001b) Chronic elevation of plasma thioredoxin: inhibition of chemotaxis and curtailment of life expectancy in AIDS. Proc. Natl. Acad. Sci. U.S.A. **98**, 2688–2693 CrossRef

85 Kang, R., Chen, R., Zhang, Q., Hou, W., Wu, S., Cao, L., Huang, J., Yu, Y., Fan, X.G., Yan, Z. et al. (2014) HMGB1 in health and disease. Mol. Aspects Med., doi:10.1016/j.mam.2014.05.001 PubMed

86 Schwiebert, E.M., Benos, D.J. and Fuller, C.M. (1998) Cystic fibrosis: a multiple exocrinopathy caused by dysfunctions in a multifunctional transport protein. Am. J. Med. **104**, 576–590 CrossRef PubMed

87 Capello, M., Ferri-Borgogno, S., Cappello, P. and Novelli, F. (2011) α-Enolase: a promising therapeutic and diagnostic tumor target. FEBS J. **278**, 1064–1074 CrossRef PubMed

88 McDonnell, C.M., Benn, D.E., Marsh, D.J., Robinson, B.G. and Zacharin, M.R. (2004) K40E: a novel succinate dehydrogenase (SDH)B mutation causing familial phaeochromocytoma and paraganglioma. Clin. Endocrinol. (Oxf.) **61**, 510–514 CrossRef PubMed

89 Alam, N.A., Olpin, S. and Leigh, I.M. (2005) Fumarate hydratase mutations and predisposition to cutaneous leiomyomas, uterine leiomyomas and renal cancer. Br. J. Dermatol. **153**, 11–17 CrossRef PubMed

90 Ahmad, S.S., Glatzle, J., Bajaeifer, K., Bühler, S., Lehmann, T., Königsrainer, I., Vollmer, J.P., Sipos, B., Ahmad, S.S., Northoff, H. et al. (2013) Phosphoglycerate kinase 1 as a promoter of metastasis in colon cancer. Int. J. Oncol. **43**, 586–590 PubMed

91 Hance, M.W., Dole, K., Gopal, U., Bohonowych, J.E., Jezierska-Drutel, A., Neumann, C.A., Liu, H., Garraway, I.P. and Isaacs, J.S. (2012) Secreted Hsp90 is a novel regulator of the epithelial to mesenchymal transition (EMT) in prostate cancer. J. Biol. Chem. **287**, 37732–37744 CrossRef PubMed

Received 6 October 2014
doi:10.1042/BST20140273