DR MARTIN RYAN

Powerful protein

While studying foot and mouth disease, **Dr Martin Ryan** discovered a new protein motif with a previously unobserved function that has led to a wide number of new directions for biomedical research

Can you explain the difference between proteins and polyproteins, and the role that 2A plays in their processing?

Proteins are composed of amino acids that are joined together, or polymerised, with peptide bonds by the ribosome. In cells, each protein is synthesised from a single, specific gene. Contrastingly, in picornaviruses such as foot and mouth disease virus, the entire complement of 14 proteins is encoded in a single gene and synthesised in a 'polyprotein' – similar to a long string of sausages that need to be cut or cleaved apart after they are made. Polyproteins are generally cleaved by proteinases, but in picornaviruses, 2A sequences occur between the upstream capsid proteins that make up the virus particle and a series of proteins downstream that are involved in the replication of the virus genome. Cleavage via 2A sequences represents an alternative method of synthesising multiple separate proteins, in which separation occurs during synthesis.

What is the innate function of 2A in picornaviruses?

The nature of the 2A protein varies between different picornaviruses. In foot and mouth disease virus, the 2A sequence is only 18 amino acids long and its primary function is to separate the capsid proteins from those involved in the viral replication. It is critical to understand how the virus grows (replicates) within an infected cell so that we can use this knowledge to combat the disease – by making new vaccines, for example.

As interest in 2A peptide sequences increases, they are being identified in an ever growing number of organisms. How does their role differ between these various systems?

Initially, we thought 2A was used only by picornaviruses, as there were no 2A-like sequences (2As) within the first genomes that were sequenced; including human, mouse, fruit fly and yeast. Thus, we were excited to find many 2As in the genome of the purple sea urchin. These occurred within two types of genetic element. The first – non-long terminal repeat retrotransposons – are found in all multicellular organisms. They come out of the genome into the cell, copy themselves and then re-insert at a new site. This is how, over evolutionary time, around 17 per cent of the human genome has come to be made up of these elements. We initially found these 2As within the genome of the sea urchin but then subsequently within a much wider range of other organisms.

The second type of genetic element is genes involved in the immune system; so, we have 2As both in virus pathogens and in the cellular defence systems that work against them. In the immunity genes 2As have been shown to have a different function; they are not found in the middle of proteins but at their beginning. This type of 2A has the additional function of targeting proteins to be secreted from the cell.

On what aspects of 2A protein sequence biology is your research currently focused?

At this point, we are focusing on two main areas. The first is trying to understand just how widespread 2As are in the genomes of multicellular organisms. The second involves the new uses of the 2A sequences that we have discovered.

Borrowing from nature

Work carried out at the **University of St Andrews**, UK, has led to the development of an efficient methodology for the co-expression of multiple proteins from a single gene using a small peptide originally identified in viruses

THE 2A OLIGOPEPTIDE sequence was initially identified in foot and mouth disease virus (FMDV), a member of the Picornaviridae family. It allows multiple discrete proteins to be synthesised from a single strand of virus RNA, which also functions as a messenger RNA (mRNA) in the infected cell. This ability to efficiently co-express multiple proteins has been a target of biological research for many years, and now the use of 2A in genetic engineering has made this possible.

Dr Martin Ryan and his group at the University of St Andrews, UK, have made consistent and groundbreaking contributions to this new and exciting field. Ryan's initial discovery of 2A in FMDV opened up a huge field of research, involving a wide range of potential biotechnological applications and research tools – the attainment of which is the main driver for his research today. Bringing these to fruition has led to significant advancements in the understanding of how 2A-like sequences (2As) operate and the nature of their evolutionary origin.

The 2A sequence was initially considered a possible site for the action of a proteinase. Two proteins could be synthesised with 2A as a site-of-cleavage linker, which would give rise to individual mature proteins when severed. Alternatively, 2A could have been a novel form of proteinase, but such enzymes are generally much longer than FMDV 2A. This led to an alternative hypothesis that the cleavage mechanism of 2A is analogous to internal ribosome entry sequences (IRES), which act as a site to initiate translation. Normally the 5' mRNA cap is required for the initiation of protein synthesis, but IRES act as a second site for recognition and initiation of translation, allowing proteins to be synthesised independently on either side.

Does the port 2A protein sequences play in the pathogenesis of viral diseases make them a potential drug target?

Just like drugs that are developed to bind the active sites of virus proteins, it would be possible to undertake drug screens for molecules that could disrupt the interactions of 2A. However, relatively few human virus pathogens encode 2A, so this would not currently be an economically viable target for drug development. Although a number of significant animal virus pathogens use 2A, for the most part drug development for diseases of domesticated animals does not represent an economically effective method of disease control, with vaccination being more cost-effective.

Has collaboration been important to the success of your research?

Elegant work on the mechanism of how 2A works was performed by Dr Jeremy Brown at Newcastle University, UK, as the senior collaborative partner. The work from his laboratory gave entirely new insights into the function of 2A. We have also collaborated with laboratories in the US and Europe on the development of 2A for gene therapy applications, and have supported research around the world by supplying plasmid clones and antibodies raised against 2A. Dr Garry Luke in my laboratory has played a key role over the years in developing this work.

RIBOSOME SKIPPING

Despite these initial theories for the mechanism of action of 2A, it has since been shown to operate in a completely different manner, termed 'ribosome skipping'. This mechanism does not involve the synthesis of a polyprotein followed by cleavage, but instead the discrete synthesis of the constituent proteins. In the case of a single strand of mRNA that encodes two proteins separated by the 2A sequence, the ribosome synthesises the first protein as normal and then continues to add the 2A sequence onto the end. Once produced, this sequence of amino acids interacts with the exit tunnel of the ribosome and prevents further elongation. To remove this blockage, the protein is released from the ribosome as though it had encountered a stop codon, and protein synthesis can resume on the mRNA downstream of the 2A. This is an example of translational control of protein biogenesis, rather than the more commonly observed transcriptional control.

Another useful aspect of the ribosome skipping mechanism, which is currently being investigated by Ryan's group, is the possibility of differentially targeting the proteins produced from a single gene. A co-translational signal sequence can be included immediately downstream of the 2A sequence in addition to the start of the gene. These two separate signal sequences will be targeted by a signal recognition particle, producing the same transport mechanism as for any other protein. This allows the downstream protein to be directed to a different location to its upstream counterpart, including a diverse range of subcellular sites or even the extracellular environment. The team has already shown that this technique can be used to direct proteins to the endoplasmic reticulum, Golgi apparatus, vacuoles, plasma membrane, as well as post-translational targeting to mitochondria. This wide range of possible sites for the resultant proteins is highly promising as it significantly adds to the utility of 2A as a co-expression system for biotechnology.

Picornaviruses are not the only species possessing a sequence that carries out this function. 2As have been found in a substantial variety of genomes, largely those of marine organisms. As the number of genomes sequenced increases ever more rapidly due to advances in sequencing technology, more and more 2As are being discovered. From this ever-expanding library of 2As, it has now become possible to carry out comparisons between sequences and attempt to determine the essential components that confer their function. A number of amino acids at specific positions in the sequences are conserved, and as such represent the 2A signature. This signature is suspected to be the region that binds to the ribosome exit tunnel and cause the skipping mechanism. Identification of the 2A signature has made the discovery of additional 2As significantly easier, as a species' genome can be systematically searched for the presence of the defining series of amino acids.

BIOTECHNOLOGICAL APPLICATIONS

Since the discovery of 2As, the St Andrews team has developed the use of these sequences into a vital genetic engineering technology for the production of transgenic organisms capable of producing a variety of proteins. This method of synthesising multiple proteins from a single gene has a huge number of potential applications. 2A technology has been shown to significantly improve protein expression efficiency in a wide variety of *in vivo* experiments. It has also simplified the process of transgenesis as well as expanding the types of genetic manipulation that can be performed. One entirely unforeseen improvement that has also arisen from the use of 2As is the increase in the genetic stability of transgenic species compared to other protein co-expression methods.

A particular problem with traditional protein co-expression techniques, which attempt to individually add a gene for each protein, is that not all the required genes may be consistently expressed in each cell. For example, if a biosynthetic pathway involving a number of separate enzymes is required, it is critical that they are all expressed in the same cell. If this is not the case, instead of transitioning from reactants to products a breakdown in the pathway could lead to a build-up of an intermediate. At best, this will have no effect on the cell, but at worst may cause cell death. The use of 2As as a gene expression technique ensures all required proteins for a given pathway will be produced by any cell that has adopted the single gene.

An additional advantage of using 2As is the ability to consistently express a number of proteins at an equal level. Using IRES to co-express proteins usually results in the upstream protein being expressed 10 times more than the downstream one. For example, equal expression is of particular importance for the biosynthesis of the heavy chain and light chain components of antibodies, which are highly valuable for their use as therapeutics.

VITAMIN A DEFICIENCY

Vitamin A deficiency is a major global health issue which affects hundreds of millions of people and is particularly prevalent in Asia and Africa. It significantly impacts younger people,

causing 500,000 children to become blind and around 2 million child deaths each year. This problem arises because rice, the main source of food in these areas, does not produce vitamin A or its precursor β-carotene, which have a number of vital functions in the body including growth.

Since 2000, to combat vitamin A deficiency, farmers have been growing a transgenic variety of rice that includes the biosynthetic pathway for production of β-carotene. The resultant grain is rich in vitamin A and referred to as 'golden rice'. However, despite this revolution, vitamin A deficiency still remains problem.

Research by Ryan has revealed a potential role for 2As in the treatment of this condition. In 2010, to enhance the content of golden rice, 2A co-expression technology was utilised to improve the efficiency of the expression of vitamin A's biosynthetic pathway. This enabled the production of rice with nine times the level of the vitamin when compared to the expression techniques previously used. The proliferation of the use of enhanced golden rice will hopefully go some way towards eradicating this substantial problem.

EVOLUTIONARY HISTORY

Aside from the potential technological applications of 2As, Ryan's research opens up a wide number of questions and avenues of investigation into the sequences themselves. The group has expanded its studies, particularly looking into the evolutionary history of the sequence. It is currently unclear how this class of oligopeptide came to be found in such a diverse range of species without being present in the majority of genomes, though the St Andrews team is making significant advances on this front.

After extensive experimentation, Ryan and colleagues have arrived at the conclusion that 2As evolved separately on multiple occasions. In one study they investigated 59 RNA viruses and concluded that 2As have emerged independently at least six times. This is a surprisingly large number from a small sample of viruses and indicates that the number of occurrence events throughout nature is likely to be significantly larger. Their research is also looking into the occurrence of horizontal gene transfer between species as a source of the diversity.

PROTEIN EXPRESSION FOR THE FUTURE

Since January 2008, over 400 patent applications regarding the use of 2As have been filed, of which 30 have been granted. Furthermore, the 2012 Nobel Prize in Physiology or Medicine went to Professor Shinya Yamanaka who used 2A to efficiently co-express four transcription factors in his work on regenerative medicine. These transcription factors were able to drive a mature, differentiated cell back to a pluripotent stem cell state, something which was previously considered impossible.

It is clear from the rapid and diverse uptake of 2A expression techniques that the advances made by Ryan and his group have made a significant and lasting impact. The team intends to continue leading the field in co-expression technology, and Ryan is optimistic about its potential future impact. "The scope of this co-expression technology is immense: one can easily envisage the incorporation of entire biochemical pathways into transgenic organisms, the development of new biosynthetic pathways, the co-expression of multivalent immunogens or the production of multi-functional virus nanoparticles, to name but a few applications," he concludes.

In 2010, to enhance the content of golden rice, 2A co-expression technology was utilised to improve the efficiency of the expression of vitamin A's biosynthetic pathway. This enabled the production of rice with nine times the levels of the vitamin when compared to the expression techniques previously used

INTELLIGENCE

2A AND 2A-LIKE SEQUENCES

OBJECTIVES

- To understand how widespread the occurrence of 2A-like sequences are in the genomes of multicellular organisms
- To develop new uses of 2A-like sequences discovered for biotechnology and medical applications

KEY COLLABORATORS

Dr Jeremy Brown,
Newcastle University, UK

FUNDING

Biotechnology and Biological Sciences Research Council

Medical Research Council

Wellcome Trust

CONTACT

Dr Martin Ryan
Professor

Biomedical Sciences Research Complex
School of Biology
University of St Andrews
North Haugh
St Andrews
Fife, KY16 9ST
UK

T +44 1334 463 403
E mdr1@st-andrews.ac.uk

www.st-andrews.ac.uk/ryanlab/Index.htm

MARTIN RYAN earned his undergraduate degree at King's College London, UK, before completing a Master's at Birkbeck College of the University of London. His PhD was in Molecular Virology at the University of Leicester, and he then went on to carry out postdoctoral research at Warwick University before moving to the Pirbright Institute to work on the foot and mouth disease virus. He has been at the University of St Andrews since 1994.

University of St Andrews