

**The Royal Society of Edinburgh
Prize Lectures - Session 2001-2002**

The Epigenetics of Disease

**The role of Phosphoinositide 3-Kinase in Development and
Disease**

PRIZE LECTURES

34th Bruce Preller Prize Lecture

Professor Adrian Bird, FRSE

13 June 2002

The Epigenetics of Disease

Over a period of 25 years Professor Adrian Bird has made many seminal contributions to our understanding of DNA methylation and the biological and medical significance of this key form of DNA modification. His many accomplishments include being the first to show the semi-conservative copying of DNA methylation and to obtain evidence that methyl-CpG mutates to TpG in DNA. This latter finding has had huge significance in terms of pathological mutations in human disease, including cancer.

Another key breakthrough was the identification of CpG islands which are found at the promoter regions of many human genes. This finding has also had a considerable impact on human genetics, helping people to identify disease and other genes, and has had direct relevance to cancer in terms of a silencing mechanism for tumour suppressor genes.

Over the past few years, Professor Bird has concentrated on identifying proteins that bind to methylated DNA. He has cloned the genes for a number of these and shown that these proteins repress transcription, often by recruiting histone deacetylase complexes. Thus providing a mechanistic understanding for the link between methylation, chromosome structure and gene regulation. Since Professor Bird isolated one of these genes, it has shown to be mutated in human Rett syndrome. Hence, whilst working on a fundamental aspect of molecular biology, Professor Bird has had a major influence on human genetics and medicine. His work has been recognised by a number of honours and prizes, including election to the Royal Society in 1989 and the award of the prestigious Louis-Jeantet Prize for Medicine in 1999.

Professor Bird is Director of the Wellcome Trust Centre for Cell Biology in Edinburgh and is a Governor of the Wellcome Trust.

“It is a great honour to be asked to give this lecture and I’d like to thank the RSE and its Fellows for the invitation.

I’m going to pitch this lecture at a level at which Charles Preller

might have had a chance of understanding what I’m talking about but I’m also going to include quite a lot of unpublished work which I hope will entertain those of you who are in the field.

A human being contains over one million million cells. Each cell has a nucleus containing the chromosomes carrying genetic information, and the actual molecule containing this information is DNA. We now know the entire sequence of the human genome – and here (*slide*) is a small section of the “*book of man*”. To extract information from such a sequence is no trivial matter. Knowing the human genome sequence is very much the end of the beginning rather than the beginning of the end. I expect there will be another 100 years spent working out exactly what it means and you shouldn't believe anybody who tells you any different.

What one can extract from that information is that there are genes arranged linearly along the DNA and they are expressed differently in different cells. For example one cell is a blood cell, another is a skin cell, and the reason they are different is because they express different genes. In one type of cell gene A and gene C are expressed to produce proteins which fold up, do their job and make the cell what it is. Yet gene B is silent. This lecture is really about how you silence genes.

This may seem like a strange topic because for many people the analogy for organisms is electronics. We all know that the way you

keep the TV off is to fail to switch it on; you don't have to continually take a decision to suppress it. But in an organism we have a soup and the machinery needed to make all the genes work is present all of the time. To prevent certain genes from working you have to do something special and this talk will look at the special things you do to genes to make sure they don't mess up the identity of a cell.

So how do some genes get shut down? One of the lessons we're learning is that there's rarely one answer to a biological question. This talk therefore looks at just one of the ways genes are shut down; this involves adding chemical groups, or put another way, placing 'knobs' on the DNA such that the whole region becomes 'walled off'. The suppressed gene effectively does not exist and in fact the DNA is glued up to such an extent that reversing the process is rather difficult. Ian Wilmut will know how difficult it is to reverse these effects by taking a nucleus from a differentiated cell and asking it to differentiate again. Cloning is so difficult partly because these kinds of mechanisms are very very stable, shutting these genes down for good and all.

DNA methylation is effectively a 'clamp' for genes. The sequence of bases in DNA that is methylated is CG. Here (*slide*) are two CGs

Review of the Session 2001-2002

and the actual modification is added to the C. However, because DNA is read in opposite directions on the two strands of the double helix then CG is paired with itself; so CG occurs in pairs on opposite strands of the double helix (*slide showing a methylated CG paired to unmethylated CG*). Epigenetics, to introduce the word, refers to processes that lead to heritable changes in the use of DNA that don't involve changing its sequence but involve processes added on top of the DNA that are nonetheless heritable.

Epigenetic mechanisms allow additional information added to the DNA to be passed on. Methylation does this in the following way. Here we have a very simple pattern (*slide*). One CG is methylated whereas the other one is not. When DNA replicates, the methylated strand is copied to provide another methylated CG on the new strand. However an unmethylated CG is copied with an unmethylated CG. Hence the pattern of methylation on the old DNA strand is copied onto the new one. Methyl groups are added to DNA by the methylase enzyme and it does this by looking at the DNA for methyl groups. If it finds a methylated CG paired to an unmethylated CG (i.e. hemimethylated DNA), it puts a methyl group onto the unmethylated strand. If it finds unmethylated CGs on both strands it does nothing.

The pattern arising from this activity can be transmitted for a very long time. Methyl groups don't stand out very much from the rest of the DNA, but they are nonetheless read and have biological impact. So where are all these methyl groups? For reasons that are not entirely clear genes are split up. If we look at where the CG sequences are found we find a high density at the beginning of genes in clusters that are not methylated (i.e. unmethylated). In fact most human genes have a cluster of CGs at their start positions. If it so happens that these CGs become methylated, then that gene is shut down long-term because methylation interferes with gene expression. Silence is long-term because, as I have just shown, when you methylate something it is copied forever.

So methylation is a way of shutting down genes in the long term. How does the signal get read to produce silencing? One way is that transcription factors are unable to bind to DNA because of the methyl groups. This is not terribly efficient because there are quite a lot of transcription factors that don't care about methylation at all. The other mechanism of repression occurs due to certain proteins that are attracted by DNA methylation. They attach to the methyl CG sites, changing the structure of

the chromatin and shutting down transcription of the gene. There are several of these proteins discovered in (or in collaboration with) our lab and we've been trying to work out what they do.

Before we explore these DNA binding proteins we should ask what does DNA methylation do biologically? We know that DNA methylation is important in the process of X chromosome inactivation. Females have two X-chromosomes and one is shut down because males only have one. Dosage of genes is very important in life and so one X chromosome is silenced and this depends very heavily upon methylation. Another area is genomic imprinting: sometimes you express the gene that comes from one of your parents but not the gene from the other. The reasons for this are obscure but we know that the mechanism involves DNA methylation. There are yet more occasions in which methylation plays a role in development, each time silencing certain genes.

Moving on to DNA methylation in disease, a theme throughout this talk, we can ask: "What can go wrong?" In theory what can go wrong is that the wrong genes can be silenced. This happens in cancer - a genetic disease in which mutations occur resulting in genes that are different in a nasty way, or non-functional. Methylation

on the other hand can render genes non-functional by shutting them down. This leads not only to cancer but also to fragile X syndrome, a common form of inherited mental retardation in humans. The primary reason why the necessary gene doesn't work is because of methylation. The second thing that can go wrong is that there appear to be genes that should be silenced but are not. Rett syndrome appears to be explained this way, as does the very rare ICF syndrome. Finally, methylation actually damages DNA. This wonderful system for shutting genes down permanently, has a cost associated with it; methylation actually causes mutations on quite a large scale. So those are the sorts of ways in which I'm going deal with the involvement of DNA methylation in disease. I'll do it via the interest in our lab (since that is what I know most about) in these methyl-binding proteins; proteins that love binding to DNA that is methylated and causing biological effects.

I shall begin by discussing MECP2 and Rett Syndrome. I showed you a cartoon of proteins binding to methylated DNA. Here I've drawn (in a totally non life-like way) a stretched out protein showing its different domains. It is a strange fact that although proteins function as 3D objects, you can stretch them out linearly and

Review of the Session 2001-2002

ascribe various functions to various bits. The first methyl binding protein we found was MECP2, ironically discovered whilst attempting to purify MECP1. MECP2 has a domain that takes it to methylated sites on DNA where it binds. Another domain we delineated is a transcriptional repression domain that shuts the DNA down when MECP2 has bound, enabling repression to take place.

I'm now going to talk about some medical aspects of this gene. First of all, (*slide*) these are nucleosomes with nucleosomal tails which are very conserved bits of the nucleosome that are there to have chemical messages written on them. We are going to concentrate on the bits of DNA that are not methylated. Here is a repressed gene that is not working and we assume the histone tails are clinging on, despite lack of evidence that they actually do. MECP2 comes along, binds, brings in the big machinery, lighting strikes, changes the chromatin, and now this gene is permanently off.

We were happily working away on this, safe in the knowledge that it had academic but not medical interest, when researchers in the USA discovered the gene that caused Rett Syndrome; it is mutated MECP2. Rett syndrome affects one in ten thousand girls. There is a period of 6-18 months

of normal development then they have a terrible crisis and regress dramatically. There is progressive encephalopathy, loss of expressive hand use and severe mental retardation with no speaking in almost all cases and no walking in many cases. Ataxia (shaking of the limbs), breathing abnormalities and digestive problems are all often present. All of the girls with this syndrome have new mutations in this gene because the girls do not reproduce, and boys with mutations in this gene die very early indeed. So every mutation is a new mutation. 50-100 children are born in the UK each year with Rett syndrome and there are thought to be altogether about three thousand girls with this condition in the UK. What you sometimes see are missense mutations where the base change results in the wrong amino acid inserted into the cluster around the methyl-binding domain. Elsewhere it is much more variable; there are all sorts of nonsense mutations, frameshifts etc.

Why does Rett syndrome affect girls? The gene is on the X-chromosome. Girls with Rett Syndrome are heterozygous, that is they have a functional MECP2 gene on one chromosome, and a non-functional one on the other X-chromosome. All cells have the same DNA, but what makes them different is random

X-chromosome inactivation. Some cells inactivate the mutated copy of the gene, so now they express the normal copy of MECP2 just like any normal cell. Other cells express the mutated version of the gene which is non-functional in some (and perhaps all) respects. It is these latter cells that cause the problem. Males only have one X chromosome, so if it has the mutation then death occurs in the first year of life. The girls are rescued by the fact that half of their cells have a functional gene, but are rescued at an enormous price.

We are interested in what MECP2 does. We were studying it to get at gene expression questions but now we had another reason for studying it: to understand this important syndrome. We made a gene knockout in the mouse. This (*slide*) shows a section of mouse brain to show MECP2 expression. In the knockout mice you can see there is no MECP2 in the brain. The brains of knockout mice are normal at the time they are weaned but later on they are not well. The mice have no MECP2 in their brains (the girls with Rett syndrome have some cells that are MECP2⁺ and some MECP2⁻). They are fine for a few weeks and then by about 6 weeks on average they have symptoms; they stop moving, have tremors, arrhythmic breathing and become very ill and die, or in practice are killed. These mice

have some of the attributes of Rett Syndrome but they are the males, the nulls. What about heterozygote females? Initially they were absolutely normal, normal size litters etc., but then it emerged that the heterozygous females did start to come down with the symptoms. By a year of age, more than 80% of the animals had acquired symptoms. Unlike the nulls they do not die, they stabilise and have parallel symptoms to those of Rett Syndrome in humans including an abnormal gait (they walk with their feet far apart) and there are other quite striking resemblances. We were doing this to get a model for Rett Syndrome as very little is known about its pathology except from post-mortems on brains.

So we have a mouse model. We have delayed onset of symptoms; remember in humans the onset of symptoms is at 6-18 months. In mice it is not that dissimilar, which is very strange as mice develop to a totally different time scale compared to that of humans – and there are some interesting deductions from that. They have neurological problems, reduced movement abnormal gait, limb clasping, and finally there is stabilisation like in the human syndrome. We can study these mice with the hypothesis that Rett Syndrome is caused by the failure of MECP2 mediated gene silenc-

Review of the Session 2001-2002

ing in the brain. For many in the field it is already accepted that MECP2 is a repressor, you take away the repressor and genes come springing on and that must be the problem. We have done a lot of experiments, as have others, but it is not easy to verify this hypothesis. However we are starting to get interesting information now which encourages us to believe the hypothesis may be right. The Rett Syndrome story is ongoing.

Now I'm going to talk about DNA methylation and cancer, due to silencing of the wrong genes. There is quite a lot of evidence from other labs that genes in cancer (that are mutated in many cases) can also be shut down by methylation. In other words they are still potentially functional but epigenetically they are silenced. Here are some examples: *P16inc* is a gene whose presence ensures that the cell cycle is restrained and the cell does not just rush into cell division, which of course cancer cells do without hesitation. There are familial melanomas for which there are mutations in this gene which predispose them to cancer. You also find this gene silenced by methylation. Another example is *MIH1* (a DNA repair gene) which is found to be mutated in familial colon cancer. In each of these cases you can find the gene mutated in these susceptible

families, or in tumours you can find it methylated.

This genetic evidence proves that these genes are important in cancer. Their absence is important in provoking cancer, and the fact that one finds them methylated and silenced implies strongly that methylation can do the same job as a mutation. They are silenced in sporadic tumours by methylation. Inactivation clearly confers a selective advantage. If you throw drugs in which remove the methylation you can re-express these genes and everything works fine. You can for example restore DNA repair just by demethylating the *MLH1* gene. Methylation appears to be a substitute for mutation. There is still controversy about this, but it is dying away in the face of a mountain of evidence. We are now running clinical trials to try to test some of the molecules that reduce methylation, to test their anti-cancer efficacy. Mutations are forever, but methylation is reversible in theory and so the drugs could also reverse the tumorigenesis. This is the hope, whether or not it will be delivered remains to be seen.

In this connection I now want to talk about *MBD2*, another repressor or gene silencer, in the context of cancer. *MBD2* is part of a huge machine that silences genes, by binding to the DNA. What happens when we remove the gene that encodes *MBD2*?

Taking away a gene to see what happens is the only way one can really find out what its protein does. What happens in this case is that the mice are more or less fine, but they don't look after their offspring in the slightest; an interesting behavioural point. I'm going to focus on the fact that if you take cells from the knockout mice and put methylated genes in artificially, then they can't repress them properly. Normally if you take a methylated gene and put it into a cell from any mammal (a mouse in this case) and for comparison take the same unmethylated gene and put it in the cell, then the methylated one hardly works. It works to about 2.5% of the level of the unmethylated one. These (slide) are two different cell lines derived from the MBD2⁻ mice. The striking result is that when genes are introduced into these cells, methylated genes are expressed at about 25% of the level of the unmethylated ones. If MBD2 is put back into the cells along with a reporter gene, then the repression is restored. If you put back a mutant version of MBD2 that does not bind methylated DNA, then you don't restore repression. To conclude, in this artificial assay, the cells from these mice (which are reasonably normal) do not repress methylated genes.

I'm now going to talk about work which is looking at real endog-

enous genes, i.e. part of the organism. We looked at T-helper cells, i.e. thymus cells that are involved in the immune system. Thymocytes can develop into two sorts of cells, TH1 or TH2, and set up immune reactions involved in allergy and antibody production. We take naive T-helper cells that express neither of the two genes, *interleukin-4 (IL4)* and *interferon-gamma* to a situation where they express *interferon-gamma* if they are TH1, or IL4 if they are TH2. This is a simple binary system. In MBD2⁻ cells, something goes wrong. Cells can be sorted by FACS (fluorescent activated cell sorting) with respect to expression of *interferon-gamma* or *IL4*; they migrate differently when expressing either gene. This enables gene expression to be visualised in a two dimensional pattern. The consequences of MBD2 deficiency for T-helper cell differentiation are that more cells express higher levels (and at an earlier stage) of the appropriate gene. In other words you are taking the lid off something that was previously repressing them, and there is significant derepression of the inappropriate gene; they're expressing the wrong gene, so the silencing of the wrong gene is defective. That is quite a dramatic effect to an immunologist.

The hypothesis, and there is more evidence to support it, is that MBD2 is keeping this gene

Review of the Session 2001-2002

switched off until the cell is ready. You might ask why the organism doesn't die since you are taking away something that reads the methylation signal? In response I take you back to the five different proteins for which methylation represses transcription; we are removing just one of them. The other four are still there. This kind of redundancy always gets in the way of interpreting these kinds of experiments.

So we have failure to repress an artificially transfected reporter and failure to properly repress an endogenous gene in these MBD2⁻ cells. So that supports our idea that what this gene does is to repress transcription.

What about cancer? I told you that methylated tumour suppressor genes are silenced by methylation. So we thought, since these mice are defective in repression we (in collaboration) would put them onto a cancer prone background. Normally mice live for a couple of years and occasionally they start to get tumours. Therefore the way in which one studies the effects of different chemicals and different genotypes on cancer is to get mice that are susceptible, i.e. left to themselves they live for a while and then start to get cancer. And then you can ask whether a particular treatment gives them more cancer more quickly or does it give them less? We found that

MBD2 removal gives them a lot less.

To the last part of my talk: DNA methylation and mutation. This is where epigenetics meets genetics. Epigenetics is non-genetic inheritance of changes in DNA function and it meets genetics because DNA methylation causes damage on a spectacular scale. If one looks at the total number of point mutations that give rise to a whole spectrum of human genetic diseases and ask how many of them are at CpG, the methylatable sequence, we find 26% of them are here. More than 25% of mutations that give human genetic diseases and also of mutations that arise within cells to give cancer are at CpG. This is a totally disproportionate number given the frequency of the dinucleotide. The problem is caused by water. Cytosine, under the insidious influence of water, gets deaminated about 100 times per cell per day. This (*slide*) amino group gets deaminated to give a carbonyl group and as a result you get uracil instead of cytosine. Uracil in DNA counts as a strange base, and does not pair properly with guanine, so there are two signals that say 'get me out of here', and there is a big machine that comes in to do that very efficiently. However, when methylated cytosine is deaminated (by the same reaction involving water) you get thymine, a conven-

tional DNA base. In this situation all you can get hold of is the mismatch, and repair turns out to be rather inefficient. The repair gives rise to a TG mismatch. T and G do not normally pair and if not repaired give rise to a TG CA mutation.

The TG CA mutation is the mutation so frequently found in human populations. If you look at worms and flies, they don't put methylation all over their genes, so they don't suffer from this. They don't live very long for other reasons but nevertheless by covering *our* genes with methylation we invite this problem. We just have to assume that the benefits of methylation somehow outweigh the disadvantage of mutability. But for a long time it has been suspected that there must be something that repairs. Some mechanism that recognises the T (within the TG) and says "wait a minute, that T is wrong, it's paired with a G in a methyl-CG context, we need to put back the C". The protein that we've been working on, MBD4, performs this role. It has no role in silencing transcription as is the function of the previous enzymes I have discussed; it is a DNA repair protein that attempts to undo the damage that methylation of DNA causes.

I will now talk about some of the recent unpublished evidence in favour of that idea. *In vitro* the

repair protein comes in and takes out the T resulting from the mutation of a methylated CG, bringing in all the subsequent repair processes necessary to restore a C. But one always has to be suspicious of what happens in the test tube; there are some classic examples where people have got the wrong ideas from simply looking at the test-tube. The best example was when the first DNA polymerase was isolated and was in fact an exonuclease that destroyed DNA, but under the conditions in the reaction it was being forced to put it together. So you have to check whether you have the right assay in your test-tube. The properties of MBD4 *in vitro* suggest that it reduces mutability in methyl-C by initiating repair of TG mismatches at methyl-CpG sites. If so, then knockout mice lacking MBD4 should display an increased frequency of CpG mutations. We took the so-called 'big blue mouse' (neither big nor blue but containing a row of bacteriophage lambda genomes), then exposed the mouse to no MBD4 for a while. The lambda DNA can then be removed and a very precise screen exists for looking at the mutations within its DNA. We looked at the CII gene and asked whether this region was methylated. Bisulphite analysis shows where the methylated CG exists. Millions of phage were analysed

Review of the Session 2001-2002

and the sorts of mutations found included GC to TA (where the G is replaced by a T), GC to AT, GT to CG etc., taking out one base and putting in one base. What you notice is that one category is higher than the others. One mutation occurs at a much higher frequency and it occurs at the CpG dinucleotide. So this category has increased dramatically in these mice as a result of taking away MBD4.

What has it done to the overall frequency of mutation? This is the fraction of all point mutations seen in the experiment at the methyl-CpG site. Normally, as I showed earlier in the human diseases, it is about 25 – 30% and it turns out to be the same in these mice. In fact in every organism that has been examined it is the same percentage. But when you take away the MBD4 gene you find that between 60-80% of all mutations are at the CpG site.

To conclude, I am showing you, once more, the mouse with intestinal tumour susceptibility. If you get more mutations without MBD4, do you also get more cancer? The answer is yes, which is what you expect because you're getting more mutation and mutations cause cancer. So the conclusion is that MBD4 suppresses CpG mutation *in vivo*. This establishes that there is a mechanism devoted to the repair

of methylated DNA involving MBD4.

To summarise, methyl-CpG-binding proteins mediate silencing, and we have discussed a few of these tonight. Silencing contributes to cancer. This is largely the work of other labs that has led to this conclusion but what I've shown you is new work suggesting that the methyl-binding proteins play a key role in this, one of which may turn out to be interesting therapeutically. Failure in silencing *may* cause Rett syndrome. I say *may* because it has yet to be proven, but that is the most likely explanation. You take away something that represses, then failure to repress could be causing the problem. Finally, DNA methylation incurs a mutagenic cost (which we knew already) and repair of some damage due to DNA methylation is carried out by MBD4. You may ask why it should be a hotspot for mutation if there is a repair mechanism. I could wave my hands and give all sorts of explanations – but I won't go into that right now.

Finally I want to acknowledge those who have been involved primarily in this work:

Jacky Guy, Catherine Millar, Xin Shen Man, Brian Hendrich, Jim Selfridge, Helen Barr.

In addition we collaborated with Alan Clarke and Owen Sanson in Cardiff on the cancer part of the

project. Peter Keightley did some statistics for us. Anne Hutchins and Steve Reiner at the University of Philadelphia did the immunological part of the project."

Questions and Answers.

How dramatically is gene expression actually changed in MBD2 null mice?

"This is an area currently being examined and it is possible to find quite a lot of differences. I showed the *IL4 / interferon-gamma* results already – I think there may a lot more of that. But I think the effect might be quite subtle for many genes."

Are mutations within the human MBD4 gene linked to cancer susceptibility?

"There is evidence that mutations in MBD4 can be found in cancer, but not nearly as frequently as one would like to implicate this gene in tumorigenesis. One of the things one needs to bear in mind is that if you mutate this gene there is not a lot more mutation as a result of losing it. It is actually quite subtle, producing a low rate of mutation which, even if it were not repaired, would be a low rate in terms of the timescale of cancer. It would take an enormous amount of time to hit a cancer gene at this unrestrained mutation rate. Removal of MLH1 produces many more mutations all over the genome in a very short space of time. Taking out MBD4 is

not a big advantage for a cancer cell as the mutation rate produced is too slow. "

Is it possible to target and silence specific genes using an exogenous source?

"Unfortunately this is impossible. It would be a good idea to target methylation to a specific gene so you could shut it down (for example shutting down genes that when activated cause cancer). The trouble is that although the effect of methylation is getting to be quite well understood, the part about what decides how something gets methylated in the first place is weakly understood. One would need to understand this much better before targeting could be possible."

Why is the timescale of onset of Rett Syndrome so similar to that seen in the knockout mice?

"Perhaps a neurone in a mouse at 37°C and a neurone in a human at 37°C would behave similarly if they both were lacking the same vital component. I think what is lost in the neurons is the ability to keep going. Neurons are formed and keep on working for years. I think these cells can do it for a while, but then are no longer able to do so. They will have a half-life, and you can imagine the half-life would not differ greatly from mouse to human. There is a view in the field that it is a neuro-developmental disorder and that

Review of the Session 2001-2002

there is some developmental moment at which something critical is meant to happen in the brain and without MECP2 you cannot do it. My argument against that is that these heterozygote mice have had 3 litters. They have been living a normal life for 6 months to a year. What developmental process are they waiting to go through? There isn't one in my opinion."

To what extent is this methylation gene family conserved across the animal kingdom?

"It looks as though it is conserved across all vertebrates which all have a lot of genomic methylation. The invertebrates do not have solidly methylated genomes as vertebrates do, and most of these proteins are missing too. One is found, and this is MBD2. Evolutionarily MBD2 is the ancestor of these genes. A relative of this is found in *Drosophila*.

To what extent is the CpG 'shortage' affected by methylation in other organisms such as invertebrates?

We looked for methylated genes in many organisms as we knew CpG was rare in invertebrates. One animal we looked at was the sea squirt, which is a chordate and has domains of methylated DNA, interspersed with unmethylated. If you look at the CpG frequency then the methylated patches all have very little, and the unmethylated ones have plenty. Presence of methylation over evolutionary time has eliminated CpG from parts of the genome. The bit that is not methylated has no mutagenic burden and so has not eliminated CpG. The reason for the intermediate state is because these genomes are partly methylated."

11th Caledonian Research Foundation Prize Lecture

Lewis C. Cantley

Professor of Cell Biology, Harvard Medical School and Chief, Division of Signal Transduction, Beth Israel Deaconess Medical Center, Harvard Institutes of Medicine.

10 April 2002 (Edinburgh), 11 April 2002 (Dundee)

***The role of Phosphoinositide 3-Kinase
in Development and Disease***

Lew Cantley is credited with the discovery of a family of enzymes, the PI 3-kinases, which have critical roles in regulating normal mammalian cell growth and movement as well as cellular responses to insulin. It is a discovery that has far-reaching medical implications for the treatment of cancer, diabetes and inflammatory diseases.

The significance of Lew Cantley's work is now widely acknowledged, most recently with his election last year to the U.S. National Academy of Sciences. But as with many stories of discovery and innovation, the early days were characterised by scepticism and, like many pioneers before him, Lew had to begin by swimming against the tide.

In his early career Lew was studying the regulation of ion transport mechanisms when he realised that many factors that stimulated growth and differentiation also affected membrane lipids called phosphoinositides, or PIs, which at the time were still only suspected to have signalling roles. The PI 3-kinase field began with two controversial papers from Lew's group published prominently in *Proceedings of the National Academy of Sciences and Nature* in 1984 and 1985. They showed the presence of a PI kinase activity associated with the Transforming gene products of Rous Sarcoma virus and Polyoma virus. These viral proteins were already known either to possess tyrosine kinase activity or to associate with tyrosine kinases of the host cell. Lew showed that the PI kinase activity was a distinct component of a protein complex that also contained and was regulated by the previously characterised tyrosine kinases. With the finding that the new lipid kinase generates a novel series of lipids phosphorylated at the 3-position of the inositol ring each of which functions as an intracellular signal, the field was set to explode and it has been expanding ever since.

Much of the excitement stems from the evident medical significance of PI 3-kinase dependent signalling. PI 3-kinases are activated by the majority of known growth factors and oncogenes. The links to cancer were strengthened still further with the discovery by Jack Dixon's group that a major

Review of the Session 2001-2002

tumour suppressor gene, PTEN, encodes a phosphatase that specifically antagonises PI 3-kinase dependent signalling by dephosphorylating its lipid products. Researchers studying insulin signalling had searched in vain over nearly half a century for the elusive second messenger of insulin action before the products of PI 3-kinase were shown to fulfil this role; drugs that mimic key elements of the PI 3-kinase signalling pathway now have great promise in the treatment of type II diabetes, the new epidemic of the western world. And the recently discovered gamma isoform of PI 3-kinase specifically regulates neutrophil migration to sites of inflammation so that inhibitors should be effective in treating inflammatory disorders.

Lew's work continues to ride the crest of this very vigorous wave. In today's lecture he is going to share with us some of the recent work of his laboratory in which the deletion of genes encoding the catalytic and regulatory subunits of PI 3-kinases in mice is revealing both anticipated and unexpected biological roles of this signalling pathway.

More than fifteen years ago we discovered the enzyme, phosphoinositide 3-kinase (PI3K) because of its co-purification with oncogenic protein-tyrosine kinases. Subsequent studies by our laboratory and others showed that activation of PI3K by various oncogenes resulted in enhanced growth and survival of cultured cells. The discoveries in the mid 1990s that lipid products of PI3K could directly bind to pleckstrin homology domains in a variety of protein kinases and regulators of low molecular weight GTP binding proteins revealed downstream signaling pathways that mediate PI3K responses. Of particular interest was the finding that the cellular homolog of the AKT retroviral oncogene product was activated by products of PI3K. The viral form of AKT had been shown to cause lymphomas in

mice. However, the relevance of the PI3K pathway for human cancers was not clear until the discovery of the PTEN tumor suppressor gene. This gene is mutated or deleted in a large number of advanced human cancers, especially in glioblastomas, melanomas and metastatic prostate cancer. The discovery by Jack Dixon's laboratory that PTEN is a lipid phosphatase that degrades the products of PI3K indicated that hyperactivation of the PI3K pathway was likely to be driving these metastatic cancers. Indeed, studies with cells in culture have shown that the PI3K-AKT pathway provides both a growth and survival signal.

To further understand the role of the PI3K pathway in development and disease, we generated mice in which the regulatory subunits of

class Ia PI3K were deleted. Three different class Ia PI3K regulatory genes exist (p85a, p85b and p55g) and these genes are partially redundant in function during development. However, deletion of both p85a and p85b results in embryonic lethality at day E12 due to a variety of defects, including bleeding, suggesting the importance of this pathway for vascular development. Surprisingly, although PI3K is required for most aspects of insulin signaling, mice with heterozygous loss of p85a or homozygous loss of p85b have increased insulin sensitivity. These studies indicate that in addition to the role of p85 subunits in regulating PI3K activity, these proteins play a negative role in insulin signaling that is not yet understood at the molecular level.

To better understand the PI3K-AKT signaling pathway, we developed a new method to identify *in vivo* substrates of the AKT protein kinase. The technique combines a bioinformatics approach for scanning the human proteome for proteins with motifs likely to be phosphorylated by this kinase, and an experimental

approach using phospho-specific antibodies that recognise proteins phosphorylated at sites that resemble this motif. With this procedure, we identified the protein product of the Tuberous Sclerosis 2 gene, tuberin as a substrate of AKT. Previous studies had shown that loss of tuberin results in increased cell proliferation as well as increased cell size and that patients with Tuberous Sclerosis have widespread hamartomas similar to those of patients carrying a defective PTEN gene (Cowden's disease). Our studies indicate that Tuberin is a negative regulator of cell growth, in part because of its ability to inhibit the mTOR-p70S6K pathway for controlling protein synthesis. Phosphorylation of tuberin by AKT turns off this function, allowing cell growth to proceed. The discovery that tuberin is a link between the PI3K pathway and the control of protein synthesis and cell growth provides a better understanding of the molecular basis of cancers and Tuberous Sclerosis, and suggests new targets for pharmaceutical intervention in these diseases.