

# mismatch2model

## MAINTAINING OUR GENETIC CODE



[www.mm2m.eu](http://www.mm2m.eu)

# MISMATCH2MODEL

## maintaining our genetic code

|   |    |
|---|----|
| Mending the bridge: mismatch repair to the rescue .....   | 3  |
| Meet the molecular matchmakers .....                      | 4  |
| Mistakes in the menders: mismatch repair and cancer ..... | 6  |
| Managing the team .....                                   | 7  |
| Modelling, mathematics and more .....                     | 8  |
| NEWS AND VIEWS  |    |
| Shedding a light on cellular switching .....              | 9  |
| It's all in the details .....                             | 10 |
| Computer calculations on cancer .....                     | 11 |

# Mending the bridge: mismatch repair to the rescue

**I**t is a classic movie sequence: the hero comes to a ravine spanned by a narrow bridge of wooden planks suspended between two ropes, as he runs across he loses his footing where a plank is missing, but manages to cling on with one hand and scrambles to safety. Almost every cell in our bodies contains a delicate double-strand like that bridge – our DNA. And if the planks aren't joined correctly there can be serious consequences.

**T**he planks that link the two strands of a DNA double helix are called base-pairs. A base on one strand needs to match its correct partner on the other strand in order to form a firm joint: an 'A' base pairs with a 'T' base and a 'G' base pairs with a 'C' base. This base-pair rule is what gives DNA its remarkable ability to be copied accurately over and over again. The two strands can be pulled apart and a new strand assembled on each of the original strands. For example, an A on one original strand triggers a T to be added to the new strand; in the same position on the other original strand there will be a T, so an A is added to the new strand. Therefore two identical double-stranded DNA 'bridges' are made from one. This process is called DNA replication and it provides a complete copy of the genetic information for each new cell that is made – like photocopying an instruction manual.

**W**e all grow by our cells dividing – one cell becomes two, two become four, and so on – and given that the more than 50 trillion cells of an adult human come from a single fertilized egg that's a lot of DNA replication. If a mistake is made during one cycle of replication – perhaps a C is inserted instead of a T – it might be copied the next time the DNA is replicated, and the new cells will end up with inaccurate genetic instructions. Some mistakes in DNA can alter the way cells behave and can cause them to form tumours. Therefore precise DNA replication is vital for maintaining normal health and development.

**S**o how are mistakes prevented? The molecular machine that makes new strands of DNA is called DNA polymerase and it has a 'proof-reading' mechanism to check that the base it has added forms a perfect match. However, occasionally even this proof-reading fails and a mistake slips through the net. Because of the importance of making an accurate copy – keeping all the planks of the bridge intact – cells, like all movie heroes, have a back-up plan. This extra safety mechanism is called DNA mismatch repair.

**T**here are two main steps in mismatch repair: first, the wobbly plank – for example, an A joined to a C instead of a T – needs to be recognized; and second, repair machinery needs to be directed to the new strand of DNA so the mistake can be corrected. Scientists have discovered some of the proteins involved in these two steps and what their jobs are, but there are still more details that need to be worked out. Slowly, scientists are building up a picture of how cells are able to keep the delicate DNA 'bridge' in good shape. This will help to improve the understanding of the causes of it going wrong and the consequences for our health ■





# Meet the molecular matchmakers

We all make mistakes. And when we do, we are usually anxious to put things right as soon as possible. But sometimes it can be difficult to know what 'right' is.

What about correcting mistakes in DNA? Because a new copy of DNA is made by copying one of the old threads of genetic information it might seem straightforward to know what is right. The old strand has the right information and any errors in the new strand – such as a 'letter' of genetic code that doesn't pair up with what is written on the old strand or a few extra letters inserted by accident – need to be rectified. Easy in principle, but when the cell's molecular repair machinery is faced with double-stranded DNA how can it tell the old and new strands apart?



Understanding the inner-workings of cells is a challenging business and so scientists often start by studying something far simpler than human beings: bacteria. Making correct copies of DNA is so fundamentally important for life that bacteria and humans have similar repair systems. Scientists have found that in bacteria a chemical label, called a methyl group, is added to each strand of DNA after it is made, but not straightaway. This means that there is a window of time when the old strand has its methyl label and the new strand does not; time for the repair machinery to distinguish the 'right' old strand from the potentially 'wrong' new strand.

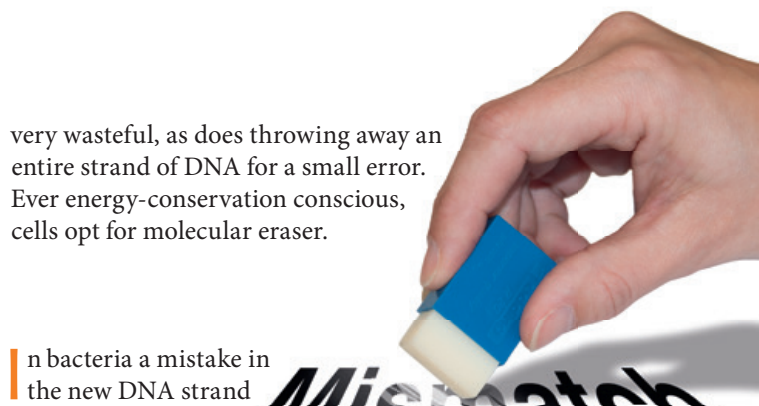


So, after you've realized that you've made a mistake and know what the right answer should be, how do you make the correction? If you make a mistake on a piece of paper you might throw that sheet in the rubbish bin and start afresh or you might just erase the mistake and insert the correction. Throwing away a sheet of paper for the sake of spelling mistake seems

very wasteful, as does throwing away an entire strand of DNA for a small error. Ever energy-conservation conscious, cells opt for molecular eraser.

In bacteria a mistake in the new DNA strand is recognized by proteins called MutS and MutL. They activate another protein, MutH, which acts like a pair of scissors to snip through the DNA near to the incorrect letter. A section of DNA, including the error, is unwound by yet another protein called UvrD, and chewed off. Then a replacement section of DNA is inserted, letter-by-letter, and the snip is sealed up again. Similarly, even if you have only got one letter wrong in a word, you might erase the complete word and write it again correctly. With the corrected word back in place, the whole sentence is seamless as though you never made the mistake.

In mammals, such as ourselves, the principles are similar to those found in bacteria. However, the details seem to be more complicated, perhaps unsurprisingly considering how much more complex a human is than a bacterium. Researchers have identified human versions of proteins such as MutS and MutL; working out the details of how mistakes are recognized, erased and corrected as each of our cells gets ready to divide is an active area of ongoing research ■



### DNA with replication error



### Mismatch recognition



### Protein complex assembly



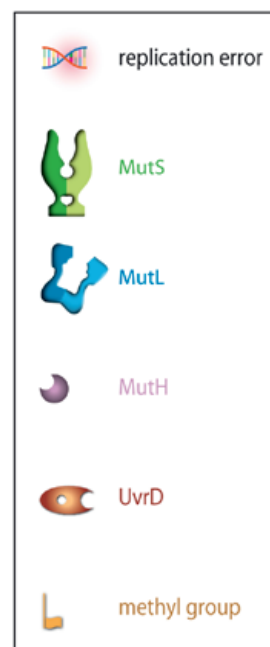
### Daughter strand nicking



### DNA unwinding



### DNA resynthesis & ligation



### The recognition and repair of replication errors in the bacterial model system *Escherichia coli*:

The replication error (for example a non-matching DNA base pair) is recognized by the protein MutS. Binding to the mismatch results in a change in the conformation of MutS and it can now interact with a second protein MutL. This complex is able to activate the nuclease MutH. MutH can distinguish which of the two DNA strands is the original one, due to the presence of a methyl group that has not yet been added to the new strand. MutH then cleaves the new strand and the helicase protein UvrD unwinds the double stranded DNA from the cleavage site towards the replication error. The unwound DNA, containing the replication error, is degraded and replaced with a new section of DNA. In this way the error is removed and the copied DNA now has the correct sequence.



# Mistakes in the menders: mismatch repair and cancer

**W**e all know that cancer could affect any of us at any time. Many of us try to decrease our risk of cancer by sticking to a healthy lifestyle, but one thing we can't do anything about is our genes. Sadly, some people are just born with a higher probability that they will get cancer. And yet there is hope; learning about the genes and cellular processes that cause a high risk of cancer can help progress toward earlier diagnosis and improved treatments.

## **W**hat does it mean to inherit a high risk of cancer?

There are some, thankfully quite rare, changes in important genes that cause a high chance of getting cancer. Hereditary Non-polyposis Colorectal Cancer (HNPCC) (also known as Lynch Syndrome) is one of these so-called cancer pre-disposition syndromes.

Colorectal cancer normally only affects people later in life, over the age of 60, but people with HNPCC are often diagnosed with this cancer in their 40s and 80% of individuals with this syndrome will get colorectal cancer during their lifetime. As well as colorectal cancer, HNPCC also causes an increased risk of tumours in other parts of the body, such as the lining of the womb (the endometrium) in women.

## **S**o what causes HNPCC?

People with HNPCC are born with a defective copy of one of several genes that are important for preventing mistakes in DNA: the DNA mismatch repair genes. You inherit two copies of all your genes: one set from your mother and one from your father. Often one defective copy of a gene can be compensated for by the fully functioning second copy, but not always. Genetic defects that cause a clinical condition when just one defective gene is inherited are called 'dominant'; HNPCC is one such condition. This means that most people with HNPCC will have a parent who also has HNPCC.

## **W**hy do problems with mismatch repair genes cause cancer?

Tumours form when a cell in the body starts to grow out of control. Normally, how often a cell divides to

form new cells is very tightly controlled and the new cells will be programmed to function correctly for the part of the body that they are in. But a tumour cell divides very rapidly and doesn't follow a programme. Because the instructions for a correctly functioning cell are written in its DNA, changes in DNA lie at the heart of cancer. Problems with the mechanism that a cell normally uses to protect its DNA from mistakes will mean that there is a much higher chance of errors in the genetic instruction manual and so a much higher chance a cell behaving abnormally will start to grow into a tumour.



## **W**hich genes are involved?

Mismatch repair is a complicated process: different proteins are needed for spotting mistakes in DNA, cutting out the incorrect DNA letters, and replacing them with a perfect genetic code. Therefore it is not surprising that there are several genes that, if they have a mistake in them, cause HNPCC. The most common cause of HNPCC is a mutation in a gene called MSH2, which is important for recognizing errors in DNA. Scientists are working to find out more about the part played by this gene and the others that can cause cancer when they don't work properly.

## **I**f mismatch repair happens in all cells, why do defects in mismatch repair genes make cancer more likely in some parts of the body?

This is a puzzle that scientists are still trying to work out. Because mismatch repair is really important when cells divide, since this is when DNA is copied and mistakes can be made, it might be that that cancer is more likely in parts of the body where cells are dividing more often. Digesting food is quite a labour-intensive business for the body and cells in the colon are continually being replaced by new ones. This might mean that the colon is particularly reliant on a fully functioning mismatch repair system.

## **H**ow does knowing about mismatch repair help?

The fact that defects in mismatch repair genes cause a very high risk of cancer highlights the importance of mismatch repair for protecting cells from becoming cancerous. Therefore, knowing more details about mismatch repair might open up new ways to treat cancer ■





# Managing the team



If you ask the managing director of a successful company what keeps their business running smoothly, a common reply will be ‘good communication and teamwork’. Scientists are finding out that in the business of mismatch repair, similar rules apply.

Good managers know the value of each of their team members, but sometimes it takes an unexpected absence to reveal which employees you really can’t do without. This is how researchers investigate the function of proteins in our bodies. In an experimental system – for example, yeast or bacteria cells that can easily be grown in a laboratory – they make a small change, known as a mutation, in the DNA code that tells the cell how to make a protein correctly. The scientists then study how important cellular processes are affected by changing that protein.

These mutation studies are teaching us a lot about the team of proteins that normally keep mismatch repair in business: each protein makes a unique contribution, if it is mutated there isn’t another than can step into its shoes. Scientists say there is virtually no ‘redundancy’ among the mismatch repair proteins.

So, if each employee has a unique role, how do all the elements of the business fit together? The key is communication. For example, in the business world the sales department, the warehouse, and the dispatch team have to talk and ensure someone is making the delivery. In the cellular world, communication is often achieved by direct contact between proteins. The 3D shape of a protein will determine which other proteins

it can interact with. For example, some proteins have intricately shaped docking sites to which only one or two specific partners can bind. Also, proteins can change the shape of these docking sites, for example after binding to a DNA mismatch. In this way a specific message can be relayed from protein to protein, in the correct order and only when necessary.

Researchers have explored the value of communication among mismatch repair proteins by making mutations that allow a protein to do its own job but prevent it from interacting with another protein, by changing a small part of its shape, for example. They found that these mutations stop the whole system from working properly, so repairing DNA really is a team effort.

Some of the mutations in mismatch repair genes that cause Hereditary non-Polyposis colorectal cancer (HNPCC) (see page 6: Mistakes in the menders) have the consequence that the proteins can no longer correctly pass the signal for ‘mismatch detected’ to the other proteins of the repair team. In cells with these mutations, DNA mismatches will remain uncorrected and will result in tumour formation. Therefore, those communication skills really are vital ■



# Modelling, mathematics and more

**T**he more we find out about biology, the more we realize how complicated it is. Or at least, that is how it can seem when you see the growing mountain of scientific data - a mountain that is increasing ever more quickly as technological improvements forge ahead.



**S**o how do scientists make sense of the large quantity of information that has been gathered using different techniques and that describes different steps in a biological process? 'Systems biology' is a rapidly growing discipline that seeks to do just that.

**T**raditionally it has been necessary for biologists to break problems down into 'bite-size' chunks. For example, although they know that a process such as DNA repair involves a whole team of proteins, for many years the analysis has been done one protein at a time. This can tell scientists what the function of its component parts are, where it goes inside cells and what makes it active or inactive. Now, a goal of systems biology is to put the pieces back together again and this is what researchers who work on mismatch repair are aiming for.



**I**mportantly, they're not using the pieces to make a static 'snapshot', but are making a dynamic model. Imagine comparing a single photograph taken at a party and video recording the whole evening. The photograph only shows you a few of the people who were there and doesn't tell you who they spoke to; the video tells you about how long the partygoers stayed for and who they interacted with.

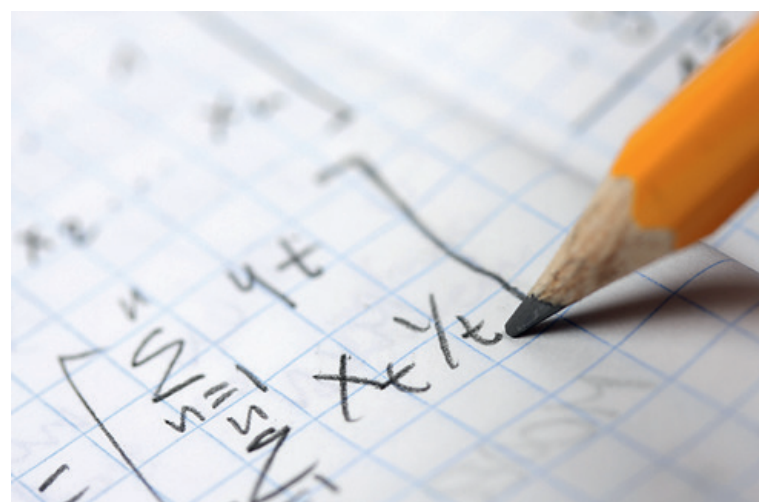
**B**ut, because we can't just video the process of mending DNA, 'seeing' the dynamics of mismatch repair is much more challenging. Mathematical modelling

is an integral part of building the dynamic picture. For example, scientists take data from experiments that show the strength of an interaction between two proteins and put that together with data about the amount of those proteins inside a cell. They can then use mathematics to make predictions about the behaviour of the proteins.

**B**ecause mismatch repair involves many proteins and many steps, putting all the information together requires a team effort from scientists with different expertise, including those who gather very detailed data from real cells and those who use computer programming to make simulations. Therefore, systems biology is not just about integrating the science; it is about integrating the scientists.

**W**hat do researchers hope to learn from making models of mismatch repair? The answer is that it can help us to understand why DNA repair can go wrong, as can happen in cancer. Imagine we're back at the party and something has gone wrong: an argument has broken out. If we rewind the video we can see that a guy started dancing with someone else's girlfriend. Likewise, if we have a dynamic model of mismatch repair we can predict what will happen if a protein changes one of its interactions.

**T**herefore, although integrating the many strands of scientific evidence is in itself a complicated process, it can help us to make sense of the true complexity of biology ■





# Shedding a light on cellular switching

Most of the time, we don't even think about it: switching on the light. Only half awake, our hand can find the button on the wall and push it. And there you go, the light is turned on. In our cells, some proteins can also be switched on and off. Not with a button of course, but by adding or removing energy in the form of a small molecule, ATP. In this way, cells control when these proteins are active or not.

In mismatch repair one of the proteins, MutS, can also be switched on. The switch in MutS, however, is not as straightforward as the light switches in our homes. When we push the button, we physically close the electrical circuit in the switch, allowing power to flow to the light bulb. And when we push the button again, the electrical circuit is broken and the light is turned off. Switching on MutS is not that black and white. Without ATP bound to it MutS is not completely off. It is still mildly active, though not sufficiently active to do its job in mismatch repair. Addition of ATP increases the activity of MutS, so it can perform its actions in the mismatch repair process.

And not only the switch is different, but also the device that is being switched on. A light bulb only has one function: emitting light. MutS however, has multiple functions. It recognizes DNA containing a mismatch and binds to it, it recruits another protein (MutL) to the site of the mismatch and afterwards it dissociates from the DNA. On top of that, it also plays a role in the release of the bound ATP, hence in its own activity switch. And all of these functions depend on the switching of MutS, though not to the same extent.

One remaining question is what triggers the switching. What in the cell is the equivalent of our hand touching the light button? By comparing the switching of MutS to that of another group of proteins, the 'mismatch2model' consortium has identified magnesium as the trigger. You might know magnesium as the white powder used by athletes, such as gymnasts and weight lifters, to improve their grip on the apparatus or lifting bar. In that case, magnesium (Mg) is in a compound together with carbonate (chalk).



In the cell, magnesium floats freely as a positively charged ion ( $Mg^{2+}$ ). In that form, it can bind to proteins. When it binds to MutS, ATP cannot bind. As a consequence, MutS is kept in the mildly active (or 'off') state. Upon removal of magnesium ATP is able to bind, turning MutS 'on'.

Strikingly, magnesium controls most, but not all functions of MutS. The recognition of the mismatch and the binding to the DNA occur in the same way whether magnesium is present or not. Obviously, the mild activity of MutS when ATP is not bound is sufficient to allow it to perform this part of its job.

The identification of magnesium as the controller of the switch in MutS is important to come to a full understanding of all details of the mismatch repair pathway. Only then researchers will be able to draw up a good model of this complicated process. Fortunately, switching on the lights does not challenge our brain to the same extent as this kind of research. That would be hard, only half awake in the morning ■



# It's all in the details

Ever asked a scientist when his research will be finished? No doubt the answer was: never! There is always more to learn. This also holds true for the 'mismatch2model' research consortium that is studying the repair of DNA mismatches. As a result of decades of meticulous experimenting, the outline of the repair process is known, the steps have been determined, the players identified and bits and pieces of the mechanism elucidated. Still the consortium wants more detail. Why? The following example will provide the answer.

Every long journey starts with a first step, as does the repair of DNA mismatches. And this first step consists of the recognition of the mistake in the DNA. When the protein MutS identifies such a mistake, it binds and bends the DNA. This also changes the shape of the protein and recently mm2m showed that MutS can then be switched on and off by two small molecules, magnesium and ATP (News & Views, September 2011). The switched-on MutS then attracts a second protein to the site, MutL. Together, these proteins initiate the repair process (a graphical representation of the complete repair pathway is on page 5 of the brochure 'Maintaining our Genetic Code').

So far, so good. But the results from different research groups started to contradict. Some groups claimed that MutS remained bound to the DNA mismatch after being switched on by ATP, while other groups claimed that MutS releases the mismatch and slides away along the DNA. Science hit deadlock. It was crucial to know which claim was correct, since MutS needs to let go of the mismatch to be able to start the actual repair reaction. Until recently, the available assays allowed researchers

to visualize either the DNA or the ATP molecules. For a conclusive answer, one needed to 'see' both.

Therefore the mm2m consortium together with collaborators developed a new approach, combining a state-of-the-art technique, called native mass spectrometry, with a newly developed mathematical algorithm. This allowed them to directly detect MutS bound to both DNA and ATP. Their results provided the answer to end the deadlock. As is often the case, both apparently contradictory claims were partly correct. Mm2m shows not one but two ATP molecules are needed to switch on MutS. The first ATP is used to 'verify' that MutS is bound to a DNA mismatch. The second ATP then causes a change in the shape of MutS, releasing the mismatch and allowing MutS to slide along the DNA.

Understanding the mismatch repair mechanism in great detail is essential for drawing up a dynamic and predictive model, the ultimate goal of the mismatch2model consortium. So if you ever think about asking a scientist when his research is finished, think again. Science will never be finished!



# Computer calculations on cancer

A dusting of snow turning to drizzle. Increasing winds and unseasonable mild temperatures. Later becoming colder – This typical weather forecast tells you to bring an umbrella and a warm scarf, just in case. How often do you actually realize that we continuously attune our daily routine to predictions based on computer models?

In science, researchers more and more depend on computer models as well. Now that we know a great deal about genes and proteins, it becomes obvious that to understand nature we must figure out how and when these molecules interlock. Depicting interactions in a graphical model can help to gain an overview. Many years of research, for example, have already yielded a detailed graphical model of the recognition and repair of replication errors called DNA mismatches (brochure 'Maintaining our Genetic Code, page 5).

But does a static model tell the whole story? As explained in the chapter 'Modelling, Mathematics and More' (brochure 'Maintaining our Genetic Code', page 8), it hardly does. Such a model is like a photograph. It is just a snapshot. You need dynamic footage, such as a video, to learn about changing interactions and to be able to quantify them. Only then, we can actually start to understand what is going on inside our cells.

So why is such a quantitative model so important? One can understand from a researcher's point of view. Curious as they are, they will always aim to unravel nature in even more detail. But a quantitative model may also be beneficiary to patients with Lynch Syndrome (also called Hereditary Non-polyposis Colorectal Cancer, HNPCC), a rare inherited disorder that gives patients a predisposition to certain types of cancer. HNPCC patients are born with a defective copy of one of the several genes involved in the repair of DNA mismatches, which increases the risk of devel-

oping colon cancer from six to about eighty percent.

A defective gene for one of the repair proteins can give rise to a lower amount of the repair protein, or a protein that cannot correctly perform its task. In some instances, it will be clear why mismatch repair is affected. For example if one of the proteins is not produced at all. Or if the graphical model predicts that two proteins can no longer form a complex. Then mismatch repair will be completely inactive. But the effect of the defective gene copy may be more subtle, for instance if a certain step in the pathway is merely slowed down. In that case the graphical model will fail to properly predict the effect on overall repair efficiency. A quantitative model that deals with the rates of the individual repair steps is then required.

Such a model is well on its way. The mismatch2model consortium has spent the last four years meticulously analyzing the kinetics of many steps in the repair process. This resulted in the very first quantitative model for mismatch repair. In the future, optimized versions of this model may well assist pathologists in assessing the severity of mismatch repair defects in tumour samples from patients. Even though these calculations will not be able to prevent or cure the disease, they may help to predict the severity of the symptoms the patient will be facing.





## Colofon

This brochure was made under the supervision of Marja Miedema, consultant Science & Society

Text: Mary Muers (Chapters)  
Marja Miedema (News & Views)

Layout by Tom de Vries Lentsch  
Images by Tom de Vries Lentsch and iStockphoto

Project coordinator mismatch2model: Joyce Lebbink  
Project manager mismatch2model: Rini de Crom

The project 'mismatch2model' is supported through Coordination Theme 1 (Health) of the Seventh Framework Programme for research and technology development (FP7) of the European Community (Grant Agreement No. HEALTH-F4-2008-223545).