

# **SFB 1309: Collaborating to Understand the Chemical Biology of Epigenetic Modifications**

---

Collaborative Research Centre 1309



# SFB 1309: COLLABORATING TO UNDERSTAND THE CHEMICAL BIOLOGY OF EPIGENETIC MODIFICATIONS

The discovery of DNA has been one of the most important findings of the last century, yet there is still much more to uncover about the ‘additional chemical layer’ brought about by the chemical modification of amino acids and nucleotide bases. A large collaboration of researchers at institutions across Germany, known as SFB 1309, is led by **Professor Thomas Carell** at the Ludwig-Maximilians University in Munich. A key focus of SFB 1309 is to elucidate the details of the second layer of information beyond the sequence layer of biochemical molecules called Watson-Crick bases within our DNA.

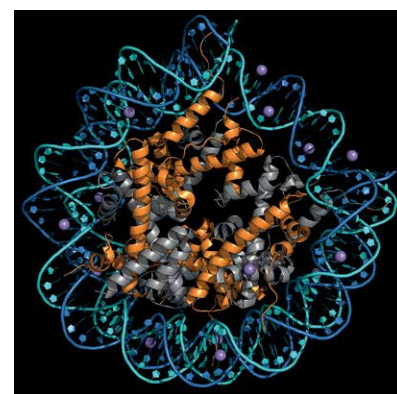
---

## Exploring the World of Epigenetic Modifications

Following the success of the Human Genome Project which ran from 1990 to 2003, genetic mutations have been the primary focus for many researchers for several decades. While the permanent alterations in the DNA sequence caused by mutations have been important in creating the diversity of species we see today, these mutations also have a negative effect on human health. The research and development of novel approaches into these permanent changes in our genetic sequence have revolutionised the way we think of diseases such as cystic fibrosis, sickle-cell anaemia, and breast cancer. However, an additional layer of information lies beyond the genetic sequence, which has captured the attention of scientists and is responsible for encoding epigenetic modification.

At the DNA level, epigenetic modifications change the way how genes are expressed. However, unlike mutations, these modifications do not bring about any change to the genetic sequence itself. These changes occur by the addition of chemical groups to the nucleotide bases. Epigenetic modifications are reversible changes and the process of epigenetics is thought to be passed down through generations. The processes of epigenetics are necessary for normal body function, but chemical modifications to the genes and proteins can also result in damage, which may initiate disease.

Epigenetic modifications are important in determining which genes are turned on or off during development – also known as gene regulation. These modifications produce a layer of information that appears to be even more complex and diverse than the sequence code. This field is largely



unexplored but has the potential to lead to the future discovery of important therapeutic targets and novel breakthrough treatments targeting various diseases.

SFB 1309 is a large-scale collaboration of researchers at institutions across Germany, dedicated to advancing knowledge in this important field. Led by Professor Thomas Carell at the Ludwig-Maximilians University in Munich, the specific aim of SFB 1309 is

## ‘These studies inform us that our genetic system is chemically more complex than thought so far.’



to elucidate the modification processes acting on nucleic acids and proteins in order to better understand the properties and functions of these new biological molecules.

SFB 1309 has brought together renowned researchers across various scientific specialities in a collaborative effort to translate the chemical language that will uncover the mystery behind epigenetic modification. To this end, the collaborative team has split the research programme into three main areas of work. Area A will investigate the structure of DNA and RNA modifications. Area B will examine the functions of the modified nucleic acids and proteins. Finally, Area C will concentrate on the development of new techniques and tools to further develop this field of work.

### Write, Read, and Erase

Epigenetic mechanisms are a normal part of human development. This is why even though cells generally contain

the same genetic instructions from our genome, the concerted expression of our genes determines the differences in the structure and function of each of the hundreds of different cell types in the human body. Epigenetic regulators determine when these genes should be expressed.

The enzymes responsible for dealing with epigenetic information are characterised into three main types, namely, ‘writer’, ‘reader’ and ‘eraser’ proteins. Epigenetic writer proteins catalyse the addition of a chemical group on to the biomolecule, also known as epigenetic mark. Epigenetic erasers are a group of enzymes responsible for removing the chemical group, and epigenetic readers usually have specific elements able to bind to the epigenetic mark, converting the chemical message into biological action.

Unfortunately, errors in epigenetic modifications may result in changes to the regulation of gene expression.



This can be hazardous in cases where a protein responsible for activating cell growth may be produced in excess, which may lead to a cell dividing far more frequently than is normal – potentially leading to the development of cancer.

Currently, there is a lack of knowledge about the process that determines which nucleic acid is targeted and the reason why the modification needs to be inserted. However, the SFB 1309 research team will explore this specific issue. Investigation of the modifications requires the unravelling of the corresponding writer, reader, and eraser proteins. ‘We will decode the chemical code above the sequence



information and we will generate molecules that can interfere with the biological processes associated with reading, writing, and erasing this code,' explains Professor Carell.

Writer, reader, and eraser proteins are at the heart of understanding epigenetic mechanisms because they play a pivotal role in triggering biological processes. The SFB 1309 research team aims to develop chemical entities that will interfere with the biological processes of these proteins. In order to understand their functional implications, the researchers must also decode the epigenetic sequence to identify the position of the chemical modifications on the DNA and RNA molecules.

Professor Carell and the SFB 1309 collaborators plan to examine the writer, reader, and eraser proteins using modern mass spectrometry techniques to identify their structure and function. Using new sequencing tools, the team aims to determine the positions of the modifications on DNA and RNA sequences. Cell biology analysis will demonstrate the biological role of the modification and examine the chemical mechanisms that induce the enzyme's ability to modify proteins and nucleic acids. The team will drive forward the development of new proteomics and sequencing methods that will also determine the position of the modified units.

### Unravelling the Layer Beyond DNA

DNA and RNA are nucleic acids consisting of a phosphate, a sugar and a base. DNA contains the sugar deoxyribose and RNA contains the sugar ribose. Five organic bases are involved in the formation of DNA and RNA – cytosine, guanine, adenine, thymidine, and uracil, which bind together to form the sequence code.

You may be asking now how the second layer of information in the genetic code was initially identified. To answer this, we must examine the previous work of researchers in the field. Scientists had initially discovered that additional DNA bases with a different chemical group existed in higher numbers, particularly in the neurons and stem cells of organisms with a high rate of reproduction.

The additional bases are 5-methyl-deoxycytidine (mdC), 5-hydroxymethyl-deoxycytidine (hmdC), 5-formyldeoxycytidine (fdC) and 5-carboxy-deoxycytidine (cadC), which are classed as modifications to the cytosine base. This discovery posed the question as to whether these four additional bases should also be considered naturally occurring DNA bases. However, at that time, detecting the chemical modifications was difficult due to the lack of technological development in the field.

Scientists later suggested that the modified cytosine bases hmdC and fdC were sites of damage to the DNA structure. However, it was discovered that an enzyme known as 'Tet' produced hmdC and fdC and these cytosine derivatives were

located in specific areas of the genome, suggesting these molecules were sequestered for a specific purpose within the cell. Scientists have shown that also a large number of proteins bind to hmdC and fdC, suggesting that there is still more to know about these modified bases than previously thought.

As Professor Carell explains, 'These studies inform us that our genetic system is chemically more complex than thought so far.' The roles of these modified DNA bases will be investigated by the SFB 1309 collaborative research team, with a view to understanding their biological function, as well as detailing the underlying mechanism of the regulation of the Tet enzymes.

The SFB 1309 team also plans to synthesise modified bases and add them into a DNA sequence to investigate how their structures change. Identification of these structural changes will help to reveal the function of the chemical modifications.

Mass spectrometry is a very sensitive analytical technique that can detect the identity of molecules in very small amounts within the organism being tested. The team plans to use mass spectrometry to identify yet undiscovered modified bases. With recent advancements in scientific technology, modern mass spectrometers are now able to detect biomolecules at much lower levels than permitted by previous methods.

### Unravelling the Layer Beyond RNA

As previously noted, the nucleotide bases within the DNA structure can undergo modifications naturally. When chemical groups alter the bases, this changes how the sequence is interpreted and, therefore, alters its properties and functions. The same occurs in the RNA sequence which codes for proteins, where post-transcriptional modifications influence how the genetic code is translated into an amino acid sequence. The range of modified nucleotide bases in RNA is far more diverse than that of DNA.

It is well established that both transfer RNA (tRNA) that is responsible for transporting amino acids, and ribosomal RNA (rRNA) that catalyses the synthesis of proteins, contain large levels of modified bases. However, scientists were surprised to find that also mRNA has considerably more modified nucleotides than they had predicted.

N<sup>6</sup>-methyl-adenosine (m<sup>6</sup>A) is the most prevalent modified nucleotide base in mRNA. A lot of attention has been paid to the modified base, which is an adenosine derivative, which was one of the first modified RNA bases to be identified. Our current understanding of m<sup>6</sup>A is that its formation involves an enzyme named methyltransferase, which acts as a writer protein, and directs itself to the adenosine base to attach a small functional group known as a methyl group. This plays an important role in regulating mRNA stability.

However,  $\alpha$ -ketoglutarate dependent enzymes such as ALKBH5 and FTO remove the methyl group, thus behaving like eraser proteins. Scientists have demonstrated that the enzymes can remove the modification from the nucleotide bases and thus suggest that m<sup>6</sup>A has a key role in cell function. The SFB 1309 research team aims to further investigate the enzymes involved. Professor Carell notes that 'It is ... speculated that m<sup>6</sup>A serves to destabilise mRNA sequences.' The m<sup>6</sup>A eraser FTO has been linked to fat mass and obesity, and m<sup>6</sup>A may be linked to nutrition and health, again demonstrating the importance of this second layer of information. Research into understanding how m<sup>6</sup>A is regulated and its role in cellular processes is currently underway.

Although many of the modified bases have been discovered, scientists are yet to uncover the biological function for the majority of these modified biomolecules. Pseudouridine is another modified nucleoside in RNA – it was first identified in tRNA and later found in mRNA as well. Sequencing data for pseudouridine in mRNA is available, however, the reader and eraser proteins have not been identified so its function is as yet unknown. Professor Carell and the collaborative team aim to identify the protein responsible for writing, editing, and erasing modifications to help reveal the function of pseudouridine in this second layer.

Scientists have also discovered relatively recently a modified base known as ac<sup>4</sup>C that is formed by an enzyme known as acetyltransferase NAT10. The modified base ac<sup>4</sup>C is a derivative of cytidine and plays a role in maintaining the stability of mRNA and to improve the success of translation. However, the reader and eraser proteins involved in the modification of this base are also yet to be uncovered, and this is a further aim for the SFB 1309 research team.

The SFB1309 research team is also in the process of using synthetic bases to investigate the structure of RNA. Similar to the investigation into DNA, these structures will be used to reveal the function of the modification of RNA bases.

### Strategies for the Future

Non-coding RNA (ncRNA) are RNA molecules that do not encode proteins but are important regulators in gene expression. The SFB 1309 research team is working to identify any modification that exists in ncRNA because of their importance in modulating gene expression and altering cell-signalling pathways. Professor Carell explains, 'It is currently thought that modified bases exist in ncRNA as well, but due to the difficulty in separating the ncRNA species from other RNA, this question is difficult to address. It is, however, now – with the new mass spectrometers that are available – the right time to investigate the chemical diversity of mRNA and ncRNA in more detail.'

In the post-genomic era, attention is being focused on the modifications that are important in cell division because understanding these complex pathways can help to develop important therapeutics in areas such as cancer. PolyADP ribose is a nucleic acid but it is also an important protein-modifying unit. PolyADP ribose forms a polymer and due to its similarity to DNA and RNA, it is involved in many different cell functions such as DNA repair.

A protein was recently discovered that binds specifically to polyADP ribose and regulates chromatin enzymes, which are important for protecting DNA from damage during cell division. The importance the modifier polyADP appears to have during cell division suggests that it may change the way cells are programmed to behave.

Professor Carell states 'Despite the fact polyADP ribose is named the third nucleic acid in contemporary biochemical textbooks, the interplay between polyADP ribosylation (the addition of ADP-ribose to a protein) and cellular plasticity awaits clarification.' To this end, the SFB 1309 research team will investigate the links between DNA damage and post-translational modifications of nucleic acids.

More research is needed to investigate the critical links between metabolism and epigenetic control. Scientists are now suggesting that the activity of mitochondria, the powerhouse of the cell, could play an important role in the control of the epigenetic modification mechanism.

In order to identify the role of the proteins that are involved in the modification of nucleic acid, the SFB 1309 research team will use inhibitors that are able to bind specifically to the modifying proteins. Inhibitors work by attaching themselves to the protein to disrupt their biological function. The team will design and synthesise these inhibitors so that they are able to observe the effect of turning off the mechanisms of these proteins.

Many human diseases are thought to have a genetic component, but the associated epigenetic mechanisms are yet to be identified. Professor Carell and his team are dedicated to uncovering critical new insights into the modified biomolecules and develop new methods to delve deeper into the phenomenon of epigenetics. As Professor Carell explains, 'We expect that understanding this layer of chemical information on biomolecules will allow us to find new avenues to treat diseases such as cancer and Angst-diseases such as schizophrenia and depression.'

In conclusion, the collaborative research conducted through SFB 1309 has the potential to provide ground-breaking insight into the true complexity of the second layer of the sequence code and open up exciting possibilities for epigenetic therapy for disease.



# Meet the researcher

**Collaborative Research Centre 1309**  
**Chemical Biology of Epigenetic Modifications**  
**Ludwig-Maximilians-Universität München**  
**München**  
**Germany**

---

The spokesperson for the Collaborative Research Centre 1309 is Professor Thomas Carell. Professor Carell received his PhD in chemistry from the University of Münster in Germany and then completed a postdoctoral fellowship in Cambridge, USA. He became a university lecturer in organic chemistry at ETH Zurich, Switzerland, in 1995. He commenced a professorship in organic chemistry at the Philipps-University in Marburg, Germany, in 2000, then took up a professorship at Ludwig-Maximilians-University München, also in Germany, in 2003. During this time, he has received many awards and prizes, such as the Gottfried Wilhelm Leibniz-Award of the Deutsche Forschungsgemeinschaft in 2004 and the Otto-Bayer-Prize from the Bayer-Schering Foundation in 2008.

## CONTACT

**E:** [Thomas.Carell@cup.uni-muenchen.de](mailto:Thomas.Carell@cup.uni-muenchen.de)

**W:** <https://www.sfb1309.de/>

## PRINCIPAL INVESTIGATORS

Professor Franz Bracher  
Professor Thomas Carell  
Professor Lena Daumann  
Professor Michael Groll  
Professor Anja Hoffmann-Röder  
Dr Eva Huber  
Dr Stefanie Kellner  
Professor Bernhard Küster  
Professor Andreas Ladurner  
Professor Kathrin Lang  
Professor Rasmus Linser  
Dr Stylianos Michalakis  
Professor Christian Ochsenfeld  
Dr Petra Rovó  
Professor Michael Sattler  
Professor Robert Schneider  
Professor Jörn Walter  
Professor Hendrik Zipse

## FUNDING

German Research Foundation DFG

