Identifying Optimum Enzymes for Gene Editing Through Simulations

Dr Rosa Di Felice



IDENTIFYING OPTIMUM ENZYMES FOR GENE EDITING THROUGH SIMULATIONS

Genome editing offers huge benefits in both healthcare and agriculture. Therefore, developing new and improved tools for editing the genomes of humans, animals and plants is a rapidly growing area of research. **Dr Rosa Di Felice** and her team from the University of Southern California are helping to further this effort by performing computational simulations that support experimentalists in their search. They are particularly interested in studying the mechanisms involved in gene editing using enzymes, to find out how their precision can be improved.



Genome Editing

The genome represents the entire collection of an organism's DNA. Animal and plant genomes contain thousands of genes, each of which ultimately determines the composition and threedimensional structure of proteins.

The synthesis of a protein in a cell is achieved through several stages. Firstly, part of the DNA code is copied onto a much smaller molecule, called RNA, which then travels to a part of the cell where proteins are made. The RNA code then determines the sequence of units called amino acids that are joined together in a chain, eventually resulting in the formation of a protein.

Genome editing, also known as gene editing, is used to make alterations in an organism's DNA, in order to change the proteins that are eventually produced. Gene editing techniques employ enzymes to insert a new gene at a specific location within the genome, or to remove one. This technology holds promise in a range of different fields, from removing disease-causing genes in humans, to incorporating genes into crops that improve their tolerance to drought or pests.

Although gene editing is now far more precise than older methods of genetic engineering, there is often still a risk that unintended edits will be made to other parts of the genome – called 'off-target edits'. Therefore, researchers are currently trying to figure out how to make genome editing techniques more accurate and precise.

Computational modelling can aid this quest, by allowing researchers to understand the gene editing mechanism in more details, so that they can figure out ways to optimise it. Dr Rosa Di Felice and her team from the University of Southern California specialise in the design and execution of these simulations of biological processes. In a recent project, the team is applying their expertise to understand how the enzymes used in gene editing technologies recognise and process DNA, towards designing more precise versions of these enzymes.



Learning from Bacterial Defences

One of the most promising gene editing techniques is known as CRISPR/Cas9. The development of this technique was inspired by a process that occurs naturally in bacterial cells. The term CRISPR refers to a specific DNA sequence in the bacterial genome, which is used by bacteria as a means of defending itself from viruses. 'In order to be employed as a therapeutic tool to correct genetic imperfections or as a tool to artificially engineer crops for large-scale use, the protocol needs to be precisely sequence-specific.'



When a virus attacks a bacterial cell, it attaches itself to the cell wall and injects its own genetic code – often DNA – to take control. In response to this, the bacterium synthesises RNA from its own genome, specifically the CRISPR sequence. The bacterial RNA then couples with an enzyme known as Cas9, which is short for 'CRISPR-associated protein 9'. The RNA strand acts as a probe and helps the enzyme identify the invading viral DNA. Once identified, the Cas9 enzyme acts as molecular scissors, cutting up the invading DNA in order to destroy it.

The CRISPR/Cas9 mechanism has now been harnessed as a genome editing tool that can cut DNA. Specific DNA sequences in an organism's genome can be targeted by changing the sequence of the CRISPR molecule. The technique is not fully precise yet, but its potential was recognised by the Nobel Prize committee in 2020. In their recent research, Dr Di Felice and her team have been performing computational modelling studies of the CRISPR/Cas9 technique. They conduct simulations by modeling individual atoms of the CRISPR RNA, the Cas9 enzyme and its analogues, and the DNA molecule that they target. 'We interrogate binding mechanisms in CRISPR systems, with the Cas9 and Cpf1 enzymes, by means of enhanced sampling molecular dynamics simulations, with the aim of proposing a structural framework to improve specificity,' explains Dr Di Felice.

Targeting the Desired DNA

DNA is made up of a pair of strands, which are wound together into the well-known double-helical structure. One strand can attach to the other because the units that make up each strand – called 'bases' – fit together like the pieces of a jigsaw puzzle. The pieces along one DNA strand can only be paired to the other if their bases are a suitable match. This pairing of bases between different strands can explain how CRISPR RNA can recognise the target DNA sequence. When the CRISPR RNA comes in contact with DNA, it can combine with it only if their bases are a good match. If there is a match, then the enzyme becomes activated. This active state is a signal that instructs the enzyme to cut the pair of DNA strands.

As the human genome is much longer and more diverse than those of viruses, there is a greater chance of matches – even 'imperfect' matches – being found in locations of the human genome that are not being targeted for editing. The same is true for the genomes of other animals and plants, which also contain thousands of genes. Even though a given sequence within a genome may not perfectly match the CRISPR RNA sequence, it can still sometimes result in DNA cleavage, leading to an off-target edit. This issue represents a major setback for CRISPR/Cas9 technology.



'In order to be employed as a therapeutic tool to correct genetic imperfections or as a tool to artificially engineer crops for largescale use, the protocol needs to be precisely sequence-specific,' explains Dr Di Felice.

One way of addressing the off-target cleavage of DNA is to ensure that the Cas9 enzyme is more sequence specific, rejecting the DNA when there is not a perfect match. Dr Di Felice believes that one way to achieve such precision will require the use of a modified – or 'mutated' – form of Cas9. To help experimentalists to design an optimum mutated form of Cas9, her team set out to simulate the interactions of multiple different mutants with DNA.

Cas9 Mutants

The team simulated the natural form of Cas9 and compared it to a series of known mutated forms. Broadly speaking, the team placed the molecules in their starting positions, in accordance with experimental data as much as possible, and then gradually heated the system to the desired room temperature. At room temperature, they then sampled motions of the atoms of CRISPR RNA, target DNA and the enzyme. This reveals the possible orientations and structures that the molecules can attain, providing the team with important insights into the mechanisms involved when the enzymes interact with the target DNA.

The matching of the CRISPR RNA strand with the target DNA sequence physically involves the incremental binding of 20 RNA bases with 20 DNA bases. This binding causes the paired DNA strands to separate. The attached DNA strand is sent along one route to be cut by the enzyme, while the other 'displaced'



DNA strand is funnelled along to a different region, which Dr Di Felice's team and others refer to as the 'positive patch' of the enzyme. As the DNA strands are separated, the negatively charged displaced strand is held in place over the positively charged patch by electrostatic interactions.

As part of their simulations, the team enveloped the test system in a salt solution to ensure it was overall electrically neutral. In addition, they could tune local neutralisation of the 'positive patch' with protein mutations, changing the chemical nature of one or few amino acids. Dr Di Felice's team found that the local neutralisation of the positive patch allowed the displaced DNA strand to move more freely in space. This makes sense since there is no longer an attraction with the negatively charged DNA strand. The finding is significant because it means that the displaced DNA strand could bump into and recombine with its other strand, displacing the RNA strand in the process and then detaching itself from the enzyme before being severed. DNA strands that were wrongly identified have a chance to remain intact.

The team noted that the mutated forms of the enzyme resulted in different spatial positionings of the regions responsible for cleaving DNA. 'Our results reveal significant structural impact of the mutations, with implications for specificity,' says Dr Di Felice.

Dr Di Felice's team hopes that their findings will provide more detailed insights into how mutations of Cas9 can affect its targeting specificity. Clearly, their work will contribute to the greater goal of refining the CRISPR/Cas9 system, towards realising the full potential of the CRISPR/Cas9 system in genome editing technology.



Meet the researcher

Dr Rosa Di Felice

Department of Physics and Astronomy University of Southern California Los Angeles, CA USA

Dr Rosa Di Felice received her PhD in Physics from the University of Rome 'Tor Vergata' in Rome, Italy in 1996. Following several postdoctoral and visiting scientist placements in Brazil, Israel and the USA, and a tenured scientist appointment in the Italian National Research Council, Dr Di Felice joined the faculty at the University of Southern California in 2013, where she is currently an Associate Professor of Physics and Astronomy and Quantitative and Computational Biology. Her research group focuses on the theoretical and computational physics of biological molecules and condensed matter. Dr Di Felice has published in diverse research areas, including protein-DNA binding, quantum computing, and materials science, and has acted as a referee for several international journals and funding bodies.

CONTACT

E: difelice@usc.edu W: https://dornsife.usc.edu/labs/compbionano

KEY COLLABORATORS

Angana Ray, University of Southern California (present address: Institute of Physics, Polish Academy of Sciences) Samprita Nandi, University of Southern California Hoda Fatemi Abhari, University of Southern California

FURTHER READING

A Ray, R Di Felice, Protein-Mutation-Induced Conformational Changes of the DNA and Nuclease Domain in CRISPR/Cas9 Systems by Molecular Dynamics Simulations, Journal of Physical Chemistry B, 2020, 124, 2168.

A Ray, R Di Felice, Molecular Simulations have Boosted Knowledge of CRISPR/Cas9: A Review, Journal of Self-Assembly and Molecular Electronics, 2019, 7, 45.

