Quantum computer solves protein puzzle

January 19 2023

The D-wave quantum annealer JUPSI of the quantum computing infrastructure JUNIQ at Forschungszentrum Jülich. Credit: Forschungszentrum Jülich / Sascha Kreklau

Physicist and code specialist Dr. Sandipan Mohanty has been working on molecular biology simulations for the world's fastest supercomputers for 20 years. Such simulations help to unravel the building blocks of life and

provide new insights into cellular machinery.

Together with researchers at Sweden's Lund University, he has now gone one step further and taken the problem of protein folding to a quantum computer. D-Wave's quantum annealer JUPSI of the quantum computer user facility JUNIQ at Forschungszentrum Jülich has more than 5,000 qubits and is the first device of this size outside North America. In an interview, Sandipan Mohanty gives insight into the pioneering work.

What was the task that you set out to do, Dr. Sandipan Mohanty?

I would say what we really achieved is to show the viability of quantum computers for non-trivial research questions in our field. Quantum computers are a quite new technology and it is not yet clear how to program them when you try to solve real scientific tasks with these new machines. For example, it is rather different from solving the problem with classical high performance computing.

More specifically, we successfully studied protein folding using a very simple model. Proteins are important building blocks of life. They fulfill a wide variety of tasks. These include, for instance, the transport of substances and cell structure. And they can only fulfill all these functions if they have a very specific form, which they achieve through a process called protein folding.

One of the many reasons why there is a lot of interest in this process is the connection between neuro-degenerative diseases like Alzheimer's or Parkinson's disease with protein mis-folding. Our hope is that quantum computers will have important advantages which will further advance our understanding of such phenomena.

Why is protein folding so computationally intensive?

Proteins are long flexible chains of [amino acids](). One fascinating property of these molecules is that a large fraction of all proteins spontaneously balls up into very specific three-dimensional shapes when you put them in a solution, think of water. So, in principle, all you need to know is the [sequence of amino acids]() that make up a protein chain. The chain then automatically knows which shape it has to fold into.

When modeling this folding process on a computer, there's a lot to try out. You can think of it like trying to calculate all the different ways you can arrange a necklace, to search for the "best" arrangement. Moreover, examining each arrangement is also computationally very expensive because of the large number of particles involved. Usually this means millions of interactions to calculate for every arrangement examined.
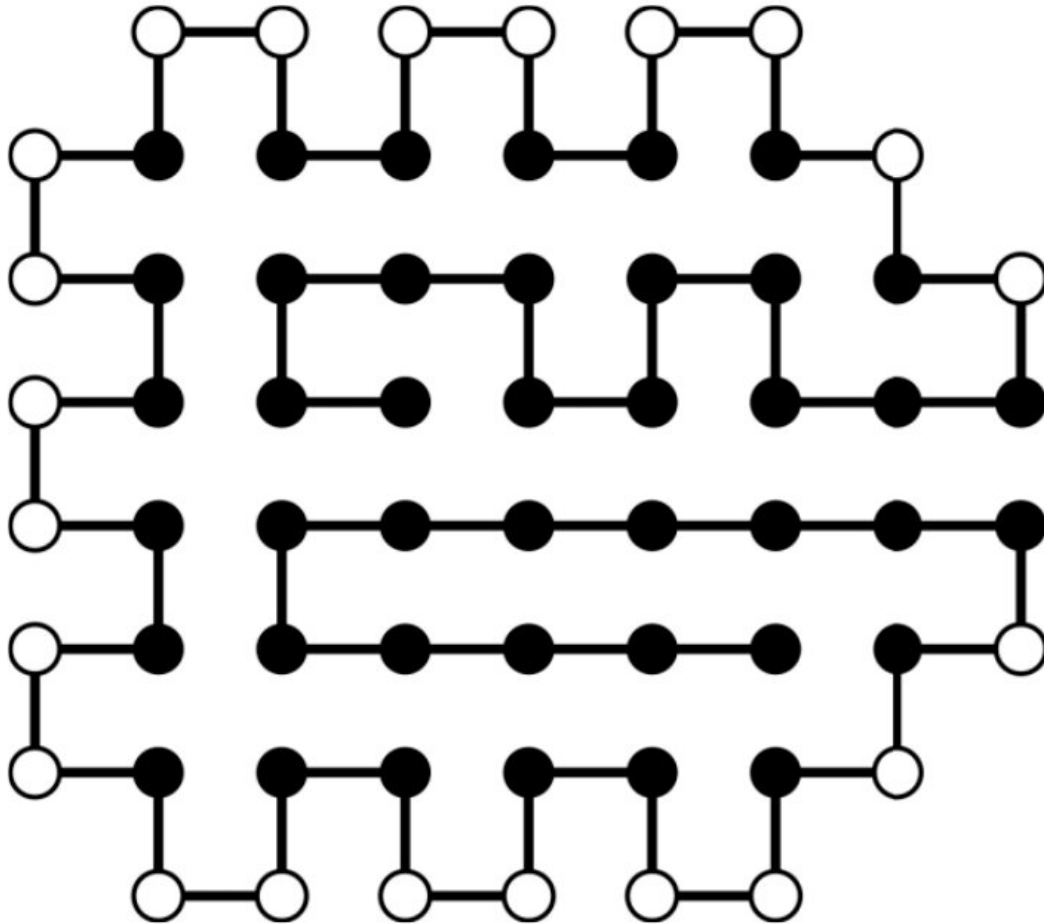
Example of a calculated structure of a chain with 64 amino acids with the lowest known energy. Filled and open symbols indicate hydrophobic (H) and water-soluble/polar (P) elements H and P, respectively. Credit: A. Irbäck et al., Phys. Rev. Research, DOI: 10.1103/PhysRevResearch.4.043013 (CC BY 4.0)

How does the quantum computer perform compared to a classical digital computer?

The task we solved is years away in complexity from the problems we

normally solve with classical supercomputers, where large scale atomically detailed simulations are common. On the D-Wave machine, we used a very reduced HP model. This simplifies the problem massively, retaining just the minimum essential physical characteristics of the folding process. We ignore the surrounding medium, divide amino acids into only two types, approximate each amino acid by a single ball which can only occupy positions on a lattice.

I should point out that as of today, even with such simplified models, being able to study chains of 64 amino acids using a quantum computer is a hard problem, which makes our results very satisfying!

Corresponding simulations can also be carried out classically. A notebook is sufficient for this. The time for the calculation does not differ much, in both cases it takes one to two minutes. However, this value is actually meaningless. Much more important is the quality of the results. And here, the quantum annealer clearly performs better.

It was rather easy to achieve 100 percent success rate in finding the lowest energy structures on JUPSI. With classical computers, on the other hand, comparable simulations only achieve 80 precent for a chain of 30 amino acids. For the more complex proteins consisting of 48 or 64 amino acid blocks, they do much worse, whereas the quantum annealer always produces the correct result here as well.

Why is the quantum computer more accurate here than a classical computer?

Because it benefits from specific aspects of the research problem. The computational effort required with classical computers, to take all relevant protein conformations into account is astronomically high. It grows exponentially with the length of the protein chain. With a chain of

two particles, there are maybe ten possibilities. With three particles, there are already a hundred. But with 100 particles—which are still rather few for a [protein](#)—one would have to calculate billions of times more variations than there are atoms in the universe.

To make any meaningful calculation at all, lots of tricks are used. Our group at the JSC and my collaborators in Sweden both specialize in so called Monte Carlo simulations. It is a procedure based on statistical physics and stochastic sampling. Although infinitely long simulations are guaranteed to produce correct results, short runs can have large errors. In practice, one tries to perform "long-enough" simulations, so that the estimated errors are acceptably small. Here lies the advantage of the quantum annealer.

This machine can, if it is programmed correctly, perform this approximation in a very direct way via its quantum mechanical couplings. Basically, it is a kind of an intricate physics experiment that automatically solves the equation. In our problem, it seems to have the effect that comparatively smaller run times are required to obtain very good answers. The fact that it works so well in practice did, however, surprise us a little.

What perspective applications does this now open up for quantum computers?

Our work marks only a first step. Most of today's quantum computers have only a few qubits. The D-Wave system has 5000, which is a lot. But for most research problems, fruitful applications of quantum computers would need even more qubits. We are still a long way from simulations such as those used in drug research done on supercomputers. I expect that we will have to wait for two or three more generations of devices to come before we will be able to solve such problems on a quantum

computer.

But I am hopeful. In contrast to the existing research we learned from, our formulation retains its simplicity with increasing system size. This opens up a possible smoother path to the study of considerably more complex problems on quantum computers.

Provided by Forschungszentrum Juelich