The quantum basis for genetic information processing

-Kapil Rajaraman (rajaramn@uiuc.edu)
PHY498 BIO, HW 1

Why does DNA have 4 nucleotides and not some other number? Why do proteins have 20 amino acids? Apoorva Patel, in a set of articles [1-4] – one of which was the focus of a cover story in the 15 April 2000 issue of the *New Scientist* [5] – claims that the answer to these questions lies in quantum dynamics. In the following essay, this work shall be overviewed and discussed. (An introduction to the topic can be found in [1], while [2] has a more technical description of the processes).

DNA and computation

Since computation is nothing but the processing of information, it is interesting to study DNA from the viewpoint of computer science. Two of the features we see in the information carried by DNA are

(a) Digitization – DNA has 4 discrete nucleotides instead of having a single continuous variable to carry information. This helps to correct errors arising from fluctuations, and also simplifies instruction sets to manipulate variables.
(b) Packing of information – Correlations (repetitive structures) in a message reduce its capacity to carry new information. Detailed analyses of DNA sequences have shown that there is little correlation between coding sequences of DNA.

Having accomplished these two requirements for efficient information transfer, it remains to find an optimal number of 'letters' in the DNA and protein alphabets. The primary criterion here is to choose as many letters as can be quickly distinguished from each other. In the case of DNA replication, whether a particular nucleotide is linked or not is the result of a simple yes/no query (i.e. the base pairing takes place or it does not). The optimization criterion is, therefore, to find the number of items that can be reliably distinguished from each other given a fixed number of queries.

Quantum search algorithms

Grover has discovered an optimal quantum database search algorithm [6] where the number of queries required Q, is related to the number of items to be distinguished N, by

$$(2Q+1)\sin^{-1}\left(1/\sqrt{N}\right)=\pi/2$$

This algorithm works by using a superposition of all possible states as the initial state. Then the amplitude of the desired state is increased iteratively (by (a) flipping the sign of the amplitude of the desired state by the query, and (b) reflecting the amplitudes about their average value) till the process is stopped at the right moment. A classical algorithm at best gives Q=O(N), whereas the quantum algorithm asymptotically goes to $Q=O(N^{0.5})$.

It is seen that an exact integral solution is Q=1, N=4. This can be used to explain base-pairing during DNA replication: a single yes/no query can distinguish between 4

possibilities. Another interesting solution is Q=3, N=20.2. Thus, we have the 3 base-pairings between the t-RNA and m-RNA transferring the genetic information to the amino acid chain. According to another paper by Patel [4] , the solution Q=2, N=10.5, is also interesting and could be indicative of the present genetic code being preceded by a more primitive one, with 2 bases coding for 10 amino acids.

How does the quantum search actually take place during DNA replication?

It is convenient to take the 4 nucleotides as the quantum basis states. As the nucleotide bases come together in random orientations, a two-step pairing process takes place which locks the nucleotides into the correct orientation. This is the equivalent of flipping the amplitude of the desired state.
After this 'push', the amplitudes then relax towards an equilibrium state like a damped pendulum. The opposite end of the oscillations around the equilibrium state is the equivalent of the reflection about average operation. At this point, the binding energy is lost to the environment, and the desired base-pairing is achieved.

But…

The biggest problem is the noisy environment within the cell, which may create quantum decoherence, which would disturb the quantum computation. While constructing a quantum computer in the lab, physicists isolate its working parts from environment by bringing temperatures down to near absolute zero. At cell temperatures, noise effects would create huge decoherence effects, which would render any quantum behavior impossible. At this point, enzymes now come into the picture.

Role of Enzymes

(a) Enzymes help to maintain the quantum coherence of the reactants – Patel claims that this is plausible, and he compares this picture with that of diamagnetic electrons, which do a good job of shielding nuclear spins from the environment, and allow nuclear spins to remain in quantum superpositions for several seconds.
(b) Enzymes are able to create superposed states of chemically distinct molecules – Nucleotides differ from each other only in terms of small chemical groups; enzymes can cut-and-paste such groups, creating a superposition of states. It should be noted that the environment is assumed to be quantum, because in a classical environment, cut-and-paste operations would result in a mixture, not a superposition of states.
This is an unusual claim, because molecules, owing to their mass, have very little quantum nature, and it would be difficult to create quantum superpositions of molecules.
Also, it is mandatory that the superposition process is only done on the growing strand, not on the intact strand. It is hypothesised that this happens by molecular bonds.

Thus, base-pairing does not occur by chance collisions (which could happen anywhere along the exposed unpaired strand), but only in the presence of enzymes.

Testing these claims

Patel has suggested a few statistical tests [3] to check the claims he has put forward. These depend heavily on being able to construct "designer DNA", which is quite possible by today's methods.

(a) Replication rates – The replication rates for classical dynamics and quantum dynamics are known. Most important, the classical rates depend on the number of nucleotides in the environment, and the quantum rates do not. In principle, therefore, it is easy to verify which dynamics is followed. However, it should be noted that disturbances from the environment will affect the quantum dynamics, making verification harder. Even in such a case, however, the quantum algorithm will work faster than its classical counterpart.
(b) Structural checks – It has been noted that the superposition of states, which is essential for the quantum algorithm to work, is done by the polymerase enzymes. This can be tested by tagging the chemical transitions by isotopic methods. In essence, a single nucleotide is tagged at locations that are replaced during transitions, as well as the groups that are not exchanged. After some time, if the two tags are still together in all the molecules, then the exchange has not taken place. If the exchange groups are found on different nucleotides, or on the polymerase enzyme, then the hypothesis is verified.

Discussion

I feel that Patel's work, though seminal, suffers from the flaw that it cannot be tested. He admits that the test on replication rates is likely to be hard due to environment effects leading to quantum decoherence. I think the test on superposition is inconclusive in that it only provides evidence of some groups being transferred between nucleotides, and not of any superposition taking place. Also, the concept of superposition arises only if the environment is quantum coherent, which brings us back in a circle to our first problem. The environment is not the only issue – even in coherent surroundings, any measurement we make on the system can modify the system if it follows quantum dynamics !
Patel's description of how the quantum search actually takes place is interesting, but it seems to be a bit contrived,in an attempt to fit natural processes to his theory.
Lastly, it is unclear how the polymerase enzymes and mRNA 'learn' how to quantum compute. This is a pretty complex multi-step process – even though Patel claims that only two quantum procedures of reflection and amplitude flipping are used – and though optimal, I think it is unlikely that replication and transcription, which have existed for billions of years, started off this way. (It is even harder to believe that a system would start off classically, and then optimize itself by shifting over to quantum processes!). This question is unanswered till the gene which codes the function of polymerase and mRNA is deciphered.

What next?

Scientists in general are taking Patel's words with more than a grain of salt – the biologists, because the process is too complicated to be realistic at a molecular level; and the physicists, because they have struggled to build a quantum computer for many years, and find it unconceivable that molecules can perform such calculations. There have been claims that the choice of 20 amino acids is not a result of optimization with respect to a quantum search algorithm. One recent example [7] explains the 64/20 genetic code redundancy by an analogy with a Bose gas with 3 particles and 4 states!
At any rate, a universally accepted and verifiable solution to the problem is not yet in sight - Apoorva Patel's work requires a significant input from the quantum computing community before the truth of his claims can be tested, and we may have to wait for some time before that happens.

References

1. A. Patel, *Why Genetic Information Processing could have a Quantum Basis*, **quant-ph/0105001**
2. A. Patel, *Quantum Algorithms and the Genetic Code*, **quant-ph/0002037**
3. A. Patel, *Testing Quantum Dynamics in Genetic Information Processing*, **quant-ph/0102034**
4. A. Patel, *Carbon – The First Frontier of Information Processing*, **quant-ph/0103017**
5. M. Buchanan, *Life force*, New Scientist (15 Apr 2000)
6. L. Grover, *A Fast Quantum Mechanical Algorithm for Database Search*, **quant-ph/9605043**
7. M.V. Altaisky and F.P. Filatov, *Genetic Information and Quantum Gas*, **quant-ph/0106123**