A publication of Ufumes.com, an online resource for bioscientists: June 2009

A future with Computational protein designing

A lot has been achieved with computer programming in the last two decades. Such feats include automation of most industrial processes and successful unmanned space exploration missions. If these much can be achieved at the binary level of programming, it's possible to achieve even more when programming is done with 20 different units if given adequate attention and intellectual reasoning.

Programming with the whole 20 amino acids gives an edge to that of computer programming since protein is a potentially tappable information-rich molecule. For instance, more than 30 million different information rich codes of polypeptides can be generated for a polypeptide sequence of 30 amino acids when constructed with the 20 amino acids as against 435 different codes that can be obtained for the same 30 units when done with 0s and 1s of the computer programming.

The recent advances in computational biochemistry have made things lots easier coupled with the successful sequencing of the entire genome of man and those of other animals. The successful crystallization and elucidation of the three dimensional structure of more than 50 000 proteins provides us with sufficient basis to understand and predict the structure and functions of proteins from known sequences.

Available to us also are lots of sophisticated sequence analytical programs such as Blast, Clawstal, Minimotive miners, EXPAsY protein analytic tools, motif scans etc., and simulation programs such as Chime, Folding@home etc. Some of these programs are available freely online. Information about DNA and protein sequences can be obtained freely from cross-referenced databases such as NCBI, EXPASY, UNIPROT, PIR, DDBJ, EMBL,etc

Computational protein designing is the use of computer software to generate chains of polypeptide sequence that would function predictably. A lot can be achieved with the computational designing of short polypeptide sequence. Since only few amino acid accounts for the catalytic function of most enzymes, it is possible to design a protein that would have two or more active sites where each site would become activated under certain conditions. The sequence of the structural part would then be designed in such a way that the final folded protein would have a 3D structure that would enable it to perform its predetermined function.

Lots of researches are already underway to advance in the discovery of proteins with novel and modified function. Most of these studies involves the modification of a particular protein's amino acid sequence to give rise to a protein whose function suites their need.

The computational designing of proteins finds potential applications in many field such as in industries; as in the designing of various analyte sensors, development of novel enzyme applicable in industrial processes. It may also be useful in pharmaceutics and medical field in the development of novel protein drugs. I envisaged a future where biochemist will design probably anything with the super informational macromolecule; Protein, by placing appropriate amino acids in a particular order.

In the design of an analyte estimator for example, it may be possible to program a polypeptide chain to undergo monitorable conformational changes in the presence of an analyte of interest. Such conformational changes may be coupled to the proportional and reversible hydrolysis of a chromorphic group attached to the polypeptide sequence.

When designing the protein rules governing the folding pattern of the proteins should be bore in mind as well the detailed knowledge of the chemistry of the individual amino acids, their bonding distances and the interaction that exist between them. All these will influence the stability and the three dimensional structure of the finally designed protein.

The polypeptide may be synthesized chemically directly, alternatively, the corresponding DNA

sequence may be chemically synthesized and then inserted into suitable vector which can then be used to amplify the expression of the designed protein.

These syntheses may pose immediate difficulty since the present technology for the chemical synthesis of polypeptide and polynucleotide sequences are expensive and could only be used to synthesize short sequences at a time.

There is need for us, biochemist to explore this emerging and interesting field. It is becoming more pertinent not only because we (Biochemist) are the only one better disposed to doing it, but more importantly because it gives us the opportunity to optimize the use of what we can freely access. We need not worry about the actual chemical synthesis for now since that can be taken care of by labs in technologically advanced country.

Even though the project might be challenging, its worth it still, because the success will finally evaluate how much will know about life.

We don't need to start or do it in a big way; all that is required is your interest, computer and the internet. It is what you can do in your spare time. The best way to start is to search for information on how each amino acid affects the secondary structure of protein. Once you have an updated information on the chemistry and interactions of the different amino acids, such knowledge can be exported into an analytical software and before you know it, you are already in real business.

We may actually be incapacitated financially, not intellectually, let's take our intellectual position in the bioscientific community.

Refer all your comments and feedback to Umeoguaju Francis, ufumeoguaju@gmail.com