Biomacromolecule Origami Technique

Qingyuan Liu
Department of Chemical and Biomolecular Engineering, University of Illinois, Urbana, Illinois 61801

This letter addresses some innovative approaches in nanobiotechnology for producing biomacromolecules with desired structures. With each approach, the capabilities and limitations will be identified. Future prospective of folding macromolecules into shape or producing shaped molecules will be stated at the end of this letter.

Nanobiotechnology has been one of the most emerging areas these recent years where nanotechnology and biotechnology has been merged together for battling with diseases, explore unknown biological field.

A DNA molecules folding technique called DNA origami technique is developed by Paul Tothemund at the California Institute of Technology there the process involves folding a long single strand of DNA by multiple smaller "staple" DNA strand [1]. By incorporating this method, numerous scientists has been able to produce desired molecules for other purpose such as drug delivery [2], molecular motor etc.

Due to its chemical nature of DNA nucleotides, only 4 of nucleotides consist all the building blocks of DNA, which makes the DNA molecule the perfect macromolecules to control it's structure but hardly able to have any chemical properties of the molecules. Therefore, molecules produced by DNA Origami technique are mostly focused on its mechanical properties but not its chemical properties. By contrast, another abundant biomacromolecules—protein is made out of 20 amino acids; with each amino acids contains different side group have different charges, it could have all kinds functions which is the reason the protein molecules are the most chemical that carried out biological functions. Plus enzymes (catalyzing protein) have structures that allow the substrate to lock into the active site of the protein, able to catalyze the reaction at an extremely fast rate. However, due to its versatility nature, and its active chemical nature, it is almost impossible to tell a protein structure just by looking at the amino acid sequences or able to alter the protein structure as protein molecules has its internal structure.

So far two general approaches would allow a protein molecules into desired shape while both them require designed peptide sequence.

The first approach is to design a long peptide sequence that can fold into the desired structure. So far only tetrahedron shape structure can be generated [3]. Known amino acids sequences which would form into α-helix structure and additionally would interact with another α-helix structured amino acid sequence is used. Every edge of the tetrahedron shape contains the two interacted α-helix amino acids. The sequence and

structure is illustrated by Figure 1a, 1b. As the sequence designed, forming into a tetrahedron would be the most energy efficient structure. This method, however, becomes exponentially more difficult as the amino acid sequence increasing or form into more complex structure. Developing another shape would require a completely different sequence. A multi-domain structure would not work in here if the whole structure were made out of a single long amino acid sequence as amino acids sequence from different domain most likely is going to interrupt the two α-helix structure's interaction. This method works great if a very small molecular structure is required.

Second technique works similar by

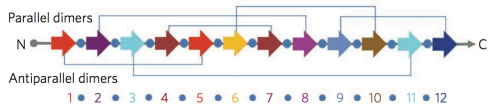

Fig 1a: representation of peptide sequence designed. 1-5; 2-8; 3-11; 4-7; 6-10; 9-12 (A-B means part A is interact with part B)

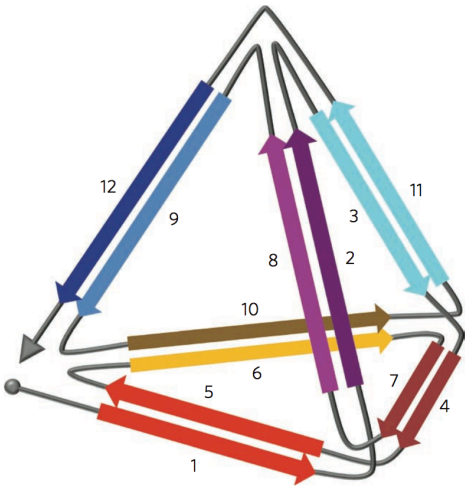

Fig 1b: The 3D structure folded by the amino acids sequence designed as illustrated in Fig 1a. 2-8, 6-10, 9-12 are parallel heterodimer. 4-7 is parallel homodimer. 1-5, 3-11 are antiparallel homodimer.

incorporating multiple peptides. The final structure depends on the amino acid interaction between the molecules not within a single molecule. A total of three types of peptides are used that all have α-helix structure [4]. One peptide can self-aggregate into a trimer but also has one active site for the other two peptides used in this experiment. The other two peptides can interact with each other or interact the first peptide mentioned. The detail can be viewed in Fig 2a, Fig 2b. This method is considered easier to operate, as the molecular structure is depends on the interaction region between two different peptides. The changing

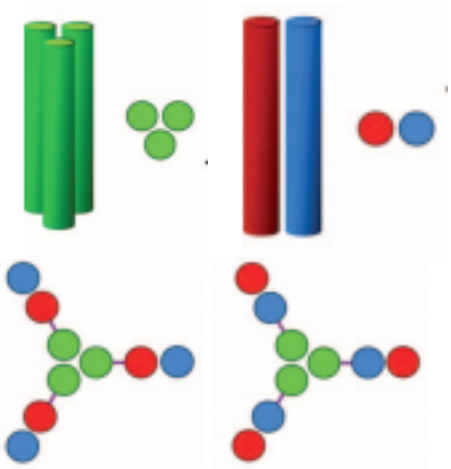

Fig 2a: Green peptide can form into a trimer with another active site could react to red or blue peptide. Red and Blue peptide can interact with each other.

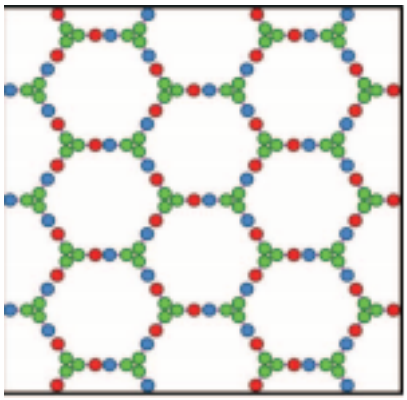

Fig 2b: A macrostructure can be formed by these three types peptide.

of the structure would only need to find peptides that would aggregate as dimer or tetramer then the whole bulking structure would be changed. However, for this method, the structure of the final product is relatively uncertain compare to the first method. Assume we require the hexagonal structure to roll into a ball shape, but the structure could end up as multiple small pieces 2D material, which did not get to roll into the ball shape.

A lot of proteins are able to interact or bind to DNA or RNA to perform their function especially those proteins function that involve replication or transcription process. A lot of known amino acid sequence has affinity to the DNA or RNA sequence. Works as the similar heuristics of the second technique, by introducing interaction between molecules would allow forming into certain desired shapes. DNA sequence can be easily folded into all kinds of structures that could produce an overall shape. Then introduce DNA sequence binding peptide structure to form into a detailed final structure. This final product has a relatively easy method for manufacturing and designing in terms of structure but having amino acid's chemical reactivity.

The major difficulty would be to determine the final structure of DNA-Protein structure and designing of this final product. Additionally, the reactivity of this final product would be extremely difficult to design or modify. Finally, this method has limitation that only amino acid sequences has affinity to DNA can be applied in this method.

The future prospect of producing desired shaped macromolecules involving amino acids would require DNA sequence designing that would allow peptide interaction to nucleotide molecules otherwise the option of the peptide selected is limited.

This area still has a lot of unsolved problems that wait to be solved.

[1]Rothemund, Paul W. K. "Folding DNA to create nanoscale shapes and patterns". Nature **440** (7082): 297–302. (2006).

[2]Garde, Damian (May 15, 2012). "DNA origami could allow for 'autonomous' delivery". fiercedrugdelivery.com. Retrieved April 2, 2016.

[3]H. Gradišar, S. Božič, T. Doles, D. Vengust, I. Hafner-Bratkovič, A. Mertelj, B. Webb, A. Šali, S. Klavžar, and R. Jerala, Nature Chemical Biology Nat Chem Biol **9**, 362 (2013).

[4]J.M. Fletcher, R.L. Harniman, F.R.H. Barnes, A.L. Boyle, A. Collins, J. Mantell, T.H. Sharp, M. Antognozzi, P.J. Booth, N. Linden, M.J. Miles, R.B. Sessions, P. Verkade, and D.N. Woolfson, Science **340**, 595 (2013).